

**Functions of SoxE in direct conversion of mouse
embryonic fibroblasts (MEFs) to neurons and neural
stem cell (NSC) maintenance**



Division of Molecular Neurobiology
National Institute of Medical Research

Department of Cell and Developmental Biology
University College London

Yuan Yang
June 2015

Thesis submitted to University College London for the degree of Master of
Philosophy

Table of contents

Abstract.....	9
----------------------	----------

Chapter 1: Introduction

1.1 Telencephalon development and neurogenesis.....	10
1.1.1 Development of embryonic telencephalon.....	10
1.1.2 Neurogenesis in mammalian cortex.....	11
1.2 bHLH proneural factors.....	14
1.2.1 Identification of proneural genes.....	14
1.2.2 Biochemical properties of proneural factors.....	14
1.2.3 Proneural factors in neural subtype specification.....	18
1.3 Epigenetic reprogramming.....	19
1.3.1 From somatic cells to pluripotent cells.....	19
1.3.2 Direct generation of neuronal cells from somatic cells.....	21
1.4 Sox family TFs in cell fate specification.....	24
1.5 Cell cycle progression and regulation.....	25

Chapter 2: Material and method

2.1 Cloning.....	28
2.2 Generation of short hairpin (sh)RNA containing plasmids.....	28
2.3 Production of lentivirus.....	29
2.4 RNA isolation and cDNA synthesis.....	29
2.5 Quantitative real time PCR.....	30
2.6 Expression microarray analysis.....	30
2.7 <i>In vitro</i> cell culture.....	31
2.7.1 P19 and human embryonic kidney (HEK) 293T cell culture.....	31
2.7.2 NS5 cell culture	31
2.7.3 Mouse embryonic fibroblasts (MEFs) isolation and culture.....	31
2.7.4 Direct reprogramming of MEF to neurons.....	32
2.8 Immunocytochemistry and image analysis.....	32
2.9 Western blot.....	33
2.10 Fluorescence-activated cell sorting (FACS) analysis.....	33

2.11 Materials.....	33
2.11.1 Oligonucleotides.....	33
2.11.1.1. Cloning.....	33
2.11.1.2. Gene expression analysis.....	34
2.11.2 Antibody and Dilution.....	35
2.11.2.1 Immunocytochemistry.....	35
2.11.2.2 Western blot.....	35

Chapter 3: Functions of Sox8 in direct conversion of mouse embryonic fibroblasts to neurons

3.1 Rational, hypothesis and aim.....	36
3.2 Re-establishment of MEFs reprogramming <i>in vitro</i>.....	37
3.3 Identification of Ascl1 downstream genes at the early stage of MEF reprogramming.....	41
3.4 Selection of Ascl1 downstream gene for further study.....	45
3.5 Sox8 is involved in Ascl1-induced MEFs reprogramming.....	47
3.5.1 Sox8 GoF reduces the generation of induced neurons (iNs).....	47
3.5.2 Sox8 LoF reduces Ascl1-mediated generation of neuronal cells from MEFs.....	50
3.6 Discussion.....	51
3.6.1 Sox8 GoF reduces the generation of MEFs iNs.....	54
3.6.2 Sox8 LoF reduces the generation of neuronal cells from MEFs.....	55
3.7 Conclusion and proposed model.....	56

Chapter 4: Functions of SoxE in neural stem cell maintenance *in vitro*

4.1 Rational, hypothesis and aim.....	57
4.2 SoxE inhibits Ascl1-induced NSC differentiation <i>in vitro</i>.....	60
4.3 SoxE activity is involved in NSC proliferation <i>in vitro</i>.....	63
4.3.1 shRNA-mediated KD efficiently reduces the expression of Sox8 and Sox9.....	63

4.3.2 SoxE is essential for NSC proliferation but not required to maintain NSC in an undifferentiated state.....	65
4.3.3 identification of SoxE downstream genes.....	67
4.3.4 Functional annotation of SoxE downstream genes.....	69
4.3.5 SoxE regulates genes involved in cell cycle regulation.....	75
4.3.6 Sox9 directly regulate cell cycle related genes.....	77
4.4 Uncovering molecular mechanisms of SoxE in regulating NSC proliferation.....	79
4.4.1 SoxE is involved in G1/S transition.....	79
4.4.2 SoxE activity is not required for mitosis.....	80
4.4.3 Cell cycle analysis using fluorescence-activated cell sorting (FACS) analysis.....	83
4.5 Regulation between SoxE and Ascl1 and within the SoxE members.....	85
4.6 Discussion.....	87
4.6.1 SoxE inhibits Ascl1-mediated NSC differentiation <i>in vitro</i>	87
4.6.2 Sox8 and Sox9 are essential and have partial functional redundancy in maintaining NSC proliferation <i>in vitro</i>	88
4.6.3 Identification and functional annotation of SoxE downstream genes.....	88
4.6.4 SoxE activity is involved in G1-phase progression and G1/S transition, but not mitosis.....	89
4.6.5 SoxE and Ascl1 in regulating NSC proliferation.....	91
Conclusions.....	92
Bibliography.....	93

List of Figures

Figure 1.1: Dorsal-Ventral Patterning of Mouse Telencephalon.....	11
Figure 1.2: Schematic of Neurogenesis and Migration in the Mouse Cortex.....	12
Figure 1.3: Classification and Structure and bHLH Proteins.....	15
Figure 1.4: Regulation of Proneural bHLH TF Activity.....	16
Figure 1.5: Generation of different cell lineage from fibroblasts using different combination of transcription factors.....	22
Figure 1.6: Different stages of cell cycle and mitosis.....	26
Figure 3.1: Direct conversion of MEFs into neurons.....	38
Figure 3.2: Identification of Ascl1 downstream genes using transcriptome-wide expression microarrays.....	42
Figure 3.3: Gene Ontology (GO) analysis of Ascl1 downstream genes.....	43
Figure 3.4: Identification of Ascl1 downstream genes involved in both NSC differentiation and MEF reprogramming.....	45
Figure 3.5: Sox8 reduces the generation of MEFs iNs.....	47
Figure 3.6: shRNA-mediated Sox8 knock down.....	50
Figure 3.7: Sox8 is required for the generation of neuronal cell from MEFs.....	51
Figure 3.8: Proposed model.....	55
Figure 4.1: Expression of Sox8, Sox9 and Ascl1 in E14.5 mouse telencephalon and graphical view of Sox8 and Sox9 protein.....	58
Figure 4.2: SoxE inhibits Ascl1-induced neuronal differentiation <i>in vitro</i>	60
Figure 4.3: shRNA-mediated SoxE knock down.....	63
Figure 4.4: SoxE activities are essential in maintaining NSC proliferation.....	65
Figure 4.5: Identification of SoxE downstream genes using genome-wide expression microarrays.....	67
Figure 4.6: Gene Ontology (GO) analysis of Ascl1 downstream genes.....	70
Figure 4.7: SoxE downstream genes are involved in cell cycle regulation.....	74
Figure 4.8: Sox9 target genes were involved in cell cycle regulation.....	77
Figure 4.9: SoxE is required for G1/S transition during cell cycle progression.....	80
Figure 4.10: SoxE activity was not involved in mitosis.....	81
Figure 4.11: Cell cycle analysis under SoxE KD conditions.....	83
Figure 4.12: Regulation between SoxE members and Ascl1.....	85

Abbreviations

Ascl1 (Mash1)	Achaete-scute complex homolog 1
AADC	Aromatic L-amino acid decarboxylase
bHLH	Basic helix-loop-helix
CNS	Central nerve system
CP	Cortical plate
CDK	Cyclin dependent kinase
D-V	Dorsal-ventral
DP	Dorsal pallium
DA	Dopaminergic
ESC	Embryonic stem cells
GoF	Gain-of-function
HMG	High-mobility-group
IPC	Intermediate progenitor cells
ICM	Inner cell mass
iPSC	Induced pluripotent stem cell
iNC	Induced neuronal cell
iMN	Induced motor neuron
KO	Knock-out
LP	Lateral pallium
LoF	Loss-of-function
MEF	Mouse embryonic fibroblast
MP	Medial pallium
MZ	Marginal zone
NSC	Neural stem cell
NCSC	Neural crest stem cells
NPC	Neural progenitor cells
PP	Preplate
PNS	Peripheral nerve system
SP	Subplate
SVZ	Sub ventricular zone
TH	Tyrosine hydroxylase

TF	Transcriptional factors
VP	Ventral pallium (VP)
VZ	Ventricle zone

Acknowledgements

I would like to thank my supervisor Dr Francois Guillemot for his support and guidance throughout the project.

I would like to give my special thank to Debbie van den Berg, who supervised me on daily bases. Thanks for sharing your knowledge and time to have discussions with me and answer my questions.

Many thanks to all the members of the lab. Thanks for all of your help and support. It was a great pleasure to work with you. More importantly, I greatly enjoyed the time we spent outside the lab. Wonderful parties at the Archway flat, thanks to Noelia and Patricia. Thanks to Koji for showing us your energetic dance moves and sharing some of your very interesting stories.

I would like to thank my thesis committee members, Vassills Pachins, Greg Elgar and Siew-Lan Ang for their support and suggestions throughout the project.

Many thanks to Eileen, who is always there to offer her help.

Abstract

Proneural transcriptional factor Ascl1 plays essential roles in inducing neuronal differentiation and subtype specification in developing central nerve system. More recently, Ascl1 has been identified as one of the critical factors that are involved in direct reprogramming of fibroblasts to functional neurons.

As Ascl1 is essential to activate neuronal specific programme in both neuronal differentiation and direct reprogramming, we hypothesis that identifying common Ascl1 downstream genes allows us to understand molecular mechanisms in both programmes.

In this study, genome-wide expression microarrays indicated that Sox8 was involved in both Ascl1-mediate reprogramming and neural stem cell maintenance *in vitro*. Both Sox8 Gain-of-Function and Loss-of-Function reduced the generation of induced neurons suggested the possibility that Sox8 was transiently expressed during the reprogramming process. Furthermore, ectopic expression of Sox8 or Sox9 inhibited Ascl1-induced neuronal differentiation of neural stem cells *in vitro*. Sox8 or/and Sox9 Loss-of-Function significantly reduced neural stem cell proliferation *in vitro*. The observation was also confirmed by genome-wide expression microarray experiments, which identified a number of cell cycle related genes were deregulated by Sox8 or/and Sox9. These data suggested that SoxE (specifically Sox8 and Sox9) activities were essential to maintain neural stem cell proliferation *in vitro*.

Chapter 1: Introduction

1.1 Telencephalon development and neurogenesis

1.1.1 Development of embryonic telencephalon

The cerebrum is the most complex region of the vertebrate central nerve system (CNS), which is raised from the embryonic structure telencephalon. The embryonic telencephalon can be divided into two regions: the dorsal region (pallium), which develops into the cerebral cortex and generates excitatory glutamatergic projection neurons; the ventral region (subpallium), which gives rise to the basal ganglia and is the birth place of inhibitory GABAergic interneurons, which later migrate dorsally into the cortex.

During embryonic development, the embryonic telencephalon can be divided into distinct progenitor domains along the dorsal-ventral (D-V) axis, which generate different parts of the adult brain (known as the D-V patterning): the medial pallium (MP) gives rise to the archicortex and the hippocampus; the dorsal pallium (DP) develops into the anlage of the neocortex; the lateral pallium (LP) generates the olfactory cortex; the ventral pallium (VP) generates the claustramygdaloid complex; the lateral (LGE) and medial (MGE) ganglionic eminences generate the striatum and pallidum (Figure 1.1, reviewed in Schuurmans and Guillemot, 2002). The initial patterning is defined by extracellular signals. The dorsal telencephalon patterning is primarily regulated by bone morphogenetic proteins (BMPs, provided by the roof plate) and Wnt (released from dorsal midline) signalling pathways (reviewed by Jessell, 2000). Later studies also suggest that zinc-finger protein GLI3 may also involved in the dorsal telencephalon patterning by repressing ventralizing Sonic hedgehog (Shh) signals in the dorsal region (Rallu et al, 2002). The ventral telencephalon patterning is regulated by the Shh, which is secreted by the floor plate (Grove et al, 1998 and Figure 1.1). At embryonic day 12.5 (E12.5), homeodomain and basic helix-loop-helix (bHLH) transcriptional factors (TFs) become highly expressed in the developing telencephalon and the cross-regulatory interactions between and within the 2 groups of TFs are essential for maintaining the progenitor identity (Figure 1.1). The expression of bHLH TF *Ascl1* (also known as *Mash1*), which is

dominantly expressed in the ventral telencephalon, is repressed by other bHLH TFs Ngns in the dorsal region (Fode et al, 2000). Furthermore, in the cortex of the *Ngn2^{KIMash1}* mice, miss-expression of *Ascl1* up-regulates ventral markers (Parras et al, 2002). These studies strongly suggest the expression of Ngns and *Ascl1* are involved in generation and maintenance of specific progenitor populations. Homeodomain TFs Pax6 and Gsh2 are two critical factors in this regulatory network. Gsh2 expression is essential for maintaining the striatal progenitor identity in vLGE. In Gsh2 null mice, expression of ventral regulators *Ascl1* and *Dlx* are lost and the ectopic expression of dorsal regulators Pax6 and Ngns can be found in the ventral region. On the other hand, Pax6 is involved in maintaining the cortical progenitor identity in the DP. In the Pax6 null mice, ectopic expression of Gsh2, *Ascl1* and *Dlx* is detected in the cortical germinal zone and Ngns are lost (Toresson et al, 2000 and Figure 1.1).

1.1.2 Neurogenesis in mammalian cortex

The adult cerebral cortex is segregated into 6 different neuronal layers (L1-L6). Each cortical layer contains distinct populations of excitatory projection neurons (originated in the dorsal telencephalon) and inhibitory interneurons (originated in the ventral telencephalon and migrated into the cortex later).

Neurogenesis started in the ventricle zone (VZ) of the dorsal telencephalon is essential for the formation of cortical layers. Starting at E10-E11, neural progenitor cells (NPCs) in the VZ undergo asymmetrical division to generate one progenitor and one neuron. The first projection neurons migrate into the preplate (PP) and form the cortical plate (CP), which subsequently develops into L2-L6. The additional CP neurons split the PP into the marginal zone (MZ) and the subplate (SP). The cortical layers are generated in an inside-out manner. Neurons generated in the SP are generated first, followed by the ones remained in L6 and L5 (deep layers) and finally the upper layers (L2, L3 and L4). The MZ develops into L1 of the cortex. (reviewed in Guillemot, 2005; reviewed in Kwan et al, 2012; Figure 1.2). The generation of GABAergic interneurons is primarily regulated by the expression of proneural gene *Ascl1* in the ventral telencephalon (see 1.2.3) and these inhibitory neurons migrate into all cortical layers in the later stage.

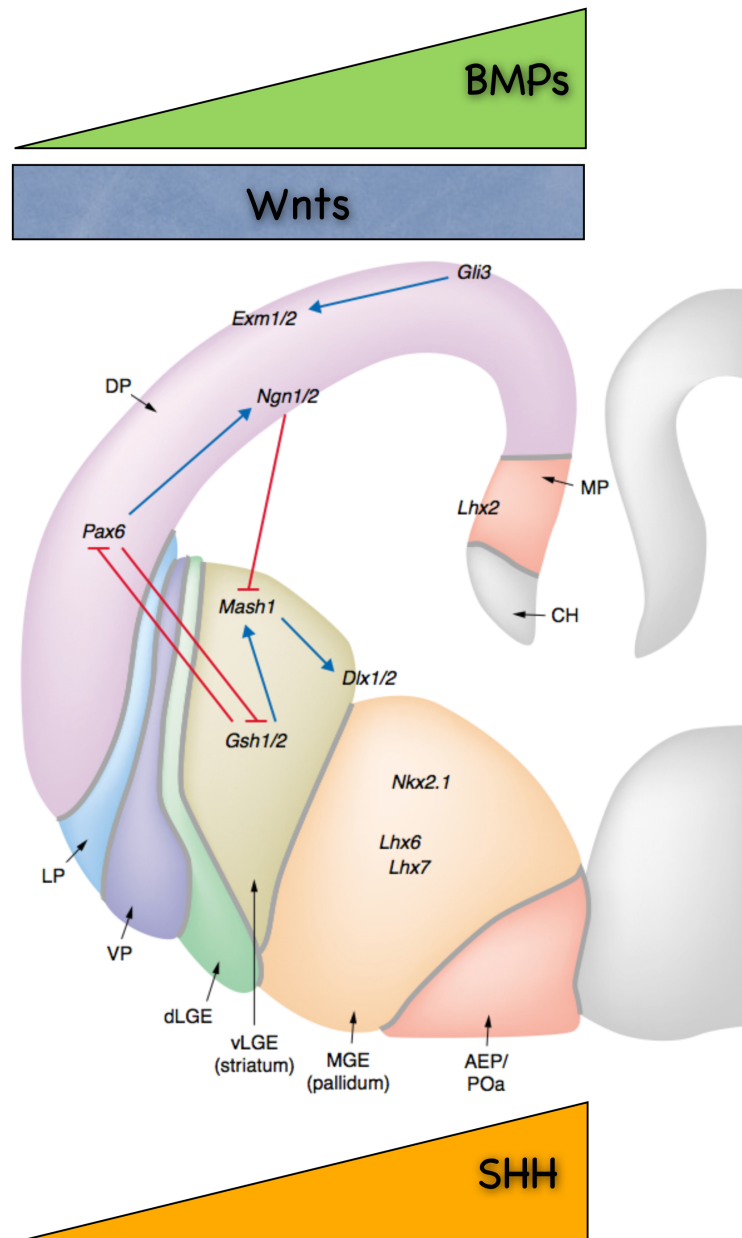


Figure 1.1: Dorsal-Ventral patterning of mouse telencephalon

Embryonic telencephalon is divided into different progenitor domains along the D-V axis. The initial D-V patterning is restricted by extracellular signalling events, including BMPs, Wnts and SHH. At E12.5, progenitors in the dorsal region express bHLH TFs *Ngn1* and *Ngn2* and homeobox genes *Emx1*, *Emx2*, *Lhx2* and *Pax6*. The ventral progenitors express bHLH TF *Mash1* (also known as *Ascl1*) and homeobox genes *Gsh1*, *Gsh2*, *Dlx1*, *Dlx2*, *Dlx5* and *Dlx6*. Arrows and T-bars refer to positive and negative controls (adapted from Schuurmans and Guillemot, 2002)

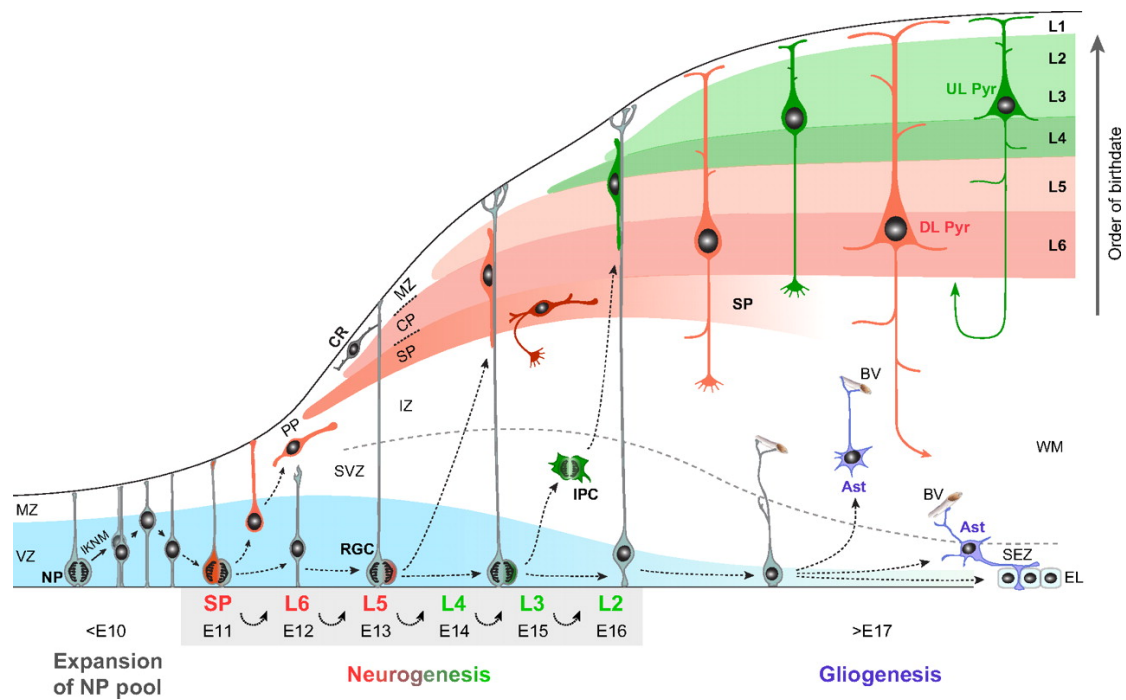


Figure 1.2: Schematic of neurogenesis and migration in the mouse cortex

Neural progenitor (NP) cells in the ventricular zone (VZ) divide asymmetrically to generate the progenitor pool (blue). Starting at E11-E12, NP cells divide asymmetrically to generate one progenitor and one neuron. The neural cells migrate into the preplate (PP), which form the cortical plate (CP) and subsequently become layers 2 to 6. Addition incoming projection neurons divide the PP into the marginal zone (MZ) and the subplate (SP). Diverse subtypes of neurons are generated. Neurons remains in the SP were generated first, followed by neurons detained in the deep layers (DL, L5 and L6) and finally the upper layers (UL, L4 to L2). The MZ eventually becomes L1. Some daughter cells of the NP cells migrate into the sub ventricular zone (SVZ) and become intermediate progenitor cells (IPCs), which divide symmetrically and contribute for the formation of upper layers.

(adapted from Kwan et al, 2012)

1.2 bHLH proneural factors

1.2.1 Identification of proneural genes

Proneural genes were first discovered in *Drosophila* in the late 1970's. In mutant flies, a set of genes were identified that were essential for sensory organs development (Garcia-Bellido, 1979). The *achete-scute* family members (*achete*, *scute*, *lethal of scute* and *asense*) and later identified *atonal* family members (*atonal*, *amos* and *cato*) share sequence similarity with each other (Gonzalez, et al, 1989; Vilares and Cabrera, 1987; Goulding et al 2000 a; Goulding et al 2000 b). Further analysis shows they also resemble the sequence of other genes, including the oncogene *myc*; the sex-determination gene *daughterless* (*da*) and muscle-determination gene *MyoD*, which lead to the identification of the bHLH DNA binding domain (Vilares and Cabrera, 1987; Murre et al 1989 a; Murre et al 1989 b; Figure 1.3 A)

The classification of the proneural functions of the *achete-scute* and the *atonal* families are founded by features shared by the two families. Firstly, in *Drosophila*, proneural genes are expressed in ectodermal cells and capable of generating neural progenitors in both peripheral nerve system (PNS) and CNS (Campuzano and Modolell, 1992). Secondly, proneural factors directly up-regulate Notch-ligands and lead to the activation of Notch signalling pathway, which is involved in balancing self-renewing and differentiation of neuronal progenitors. In mutant flies, disruption of the Notch pathway leads to excessive neuronal differentiation (Artavanis-Tsakonas et al, 1995). In the embryonic telencephalon, Notch signalling targets Hes1, Hes5 and related Hey genes inhibit the expression of *Ascl1* and *Ngns* that promote neuronal differentiation (reviewed in Pierfelice et al, 2011).

1.2.2 Biochemical properties of proneural factors

Proneural proteins are basic helix-loop-helix (bHLH) TFs. The 60 amino acid long DNA-binding bHLH motif is evolutionary conserved and contains a basic domain followed by two α -helices that are connected by a flexible loop (Figure 1.3 B). The proneural bHLH TFs are normally function as homo- or heterodimers, which is essential for its association with specific DNA sequence known as the E-Box (CANNTG) (reviewed in Massari and Murre, 2000).

Proneural bHLH TFs belong to the class B bHLH factors, which have a strict expression pattern. The expression of proneural TFs *Ascl1* and *Ngns*, for instance, is restricted to LGE and DP in the developing telencephalon (Figure 1.1). E-proteins (namely E12 and E47, products of the *E2A* gene due to alternative splicing) belong to the class A bHLH factors, which are primarily served as binding partners for proneural factors (Murre et al, 1989) (Figure 1.4 A). Proneural factors can be negatively regulated by class C repressive bHLH factors, including Id-proteins and Hes-proteins. Id-proteins inhibit proneural factor activity by forming low-stability complexes with E-proteins. The dimer is unable to bind to the DNA due to the lack of the basic domains on Id-proteins (Duncan et al 1992) (Figure 1.4 B). Hes-proteins dimerise and bind to N-box sequence (CACNAG), which recruits transcriptional repressor complex to inhibit proneural factor-induced gene transcription (Van Doren et al, 1994). Furthermore, Hes-proteins can also repress proneural factor activity by direct binding with proneural heterodimer and recruiting transcriptional repressor complex (Davis and Turner, 2001) (Figure 1.4 C and D).

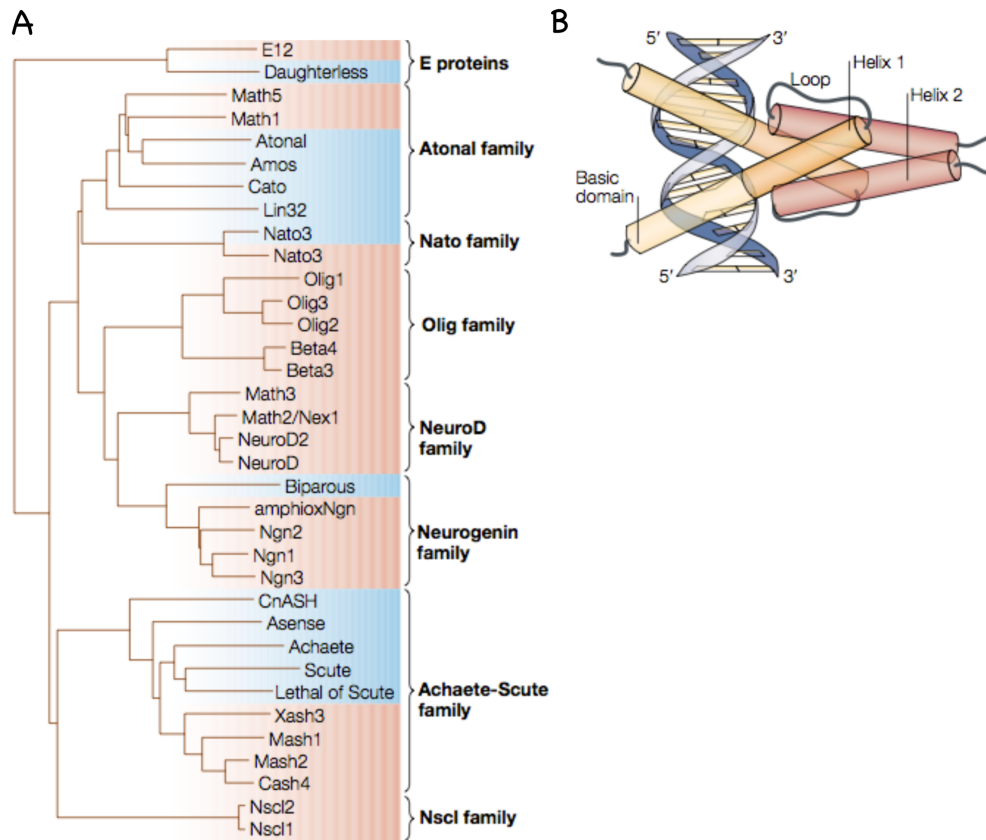


Figure 1.3: Classification and structure and bHLH proteins

A: the dendrogram of vertebrate (red) and invertebrate (blue) bHLH factors.

B: the schematic representation of the structure of bHLH dimer complexed with DNA.

(adapted from Bertrand et al, 2002)

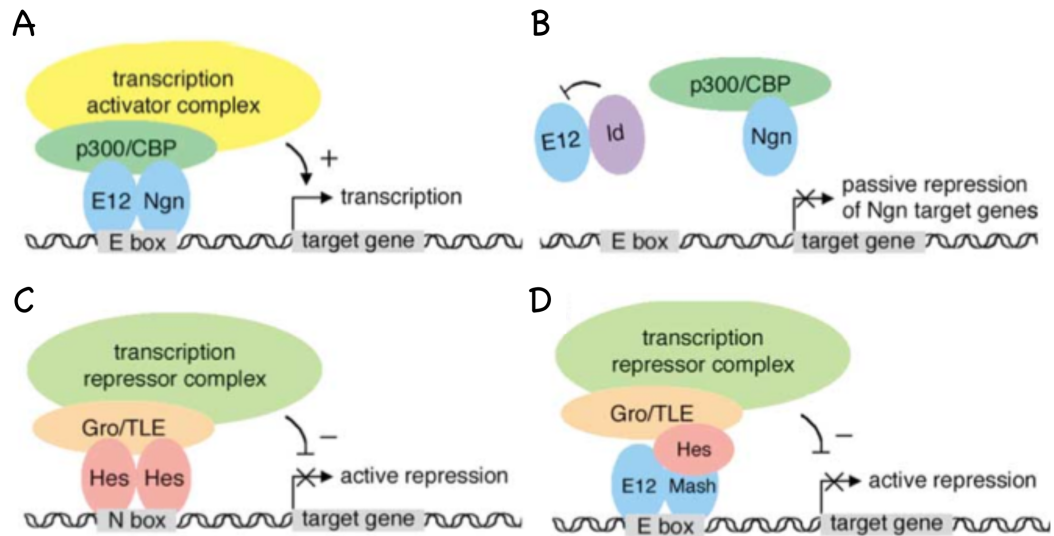


Figure 1.4: Regulation of proneural bHLH TF activity

A: bHLH TFs from heterodimers with E-proteins. The binding of the dimer with E-box recruits transcriptional activator complex and lead to the activation of target genes.

B: Id-proteins repress bHLH TF activity by forming dimers with E-proteins and preventing the formation of the proneural dimer.

C: Hes-proteins bind to the N-box sequence in the promoter region and recruit transcription repressor complex to inhibit the proneural TF activity.

D: Hes-proteins bind to proneural dimer and recruit transcription repressor complex to repress the proneural TF activity

(adapted from Ross et al, 2003)

1.2.3 Proneural factors in neural subtype specification

Proneural bHLH TFs have critical roles in neural subtype specification. In the developing mammalian telencephalon, Ngn1, Ngn2 and Ascl1 are essential TFs involved in generating distinct progenitor populations: the cortex and basal ganglia. In the developing CNS, Ngns and Ascl1 induce neural stem cell (NSCs) differentiation and give rise to mature neurons in a subtype specific manner (reviewed in Bertrand et al, 2002; Reviewed in Ross et al, 2003).

Gain-of-function (GoF) studies have demonstrated that proneural genes are sufficient to induce the formation of mature neurons. *In vitro* experiments have shown that ectopic expression of Ascl1 and Ngn2 is able to induce autonomic and sensory neurogenesis in the dissociated neural tube cultures in a BMP2 dependent manner (Lo, et al, 2002). Furthermore, upon the expression of Ascl1, adult NSCs can generate GABAergic interneurons (Berninger et al, 2007). It also has been shown that co-expression of Oligo2 and Ngn2 is able to induce motor neurons in the developing ventral neuronal tube (Mizuguchi et al, 2001) and mis-expression of Ngn2 in the ventral telencephalic progenitors can induce the formation of glutamatergic neurons (Matter et al, 2008). These data clearly suggest that proneural genes are sufficient to activate specific neuronal differentiation programs.

Loss-of-function (LoF) studies have provided further evidence to support that proneural factors are involved in neuronal identity specification. For example, in Ascl1 null mice, the ventral telencephalon, which gives rise to GABAergic interneurons has several severe defects in neurogenesis, including the missing of MGE and the reduction of LGE in size (Guillemot et al, 1993; Casarosa et al 1999). These phenotypes clearly suggest that Ascl1 is essential in the specification and/or maintenance of the ventral progenitor population. However, the possibility that the defects may due to the loss of the progenitor populations is undeniable. A series transgenic mouse models have been generated to assess the ability of Ascl1 and Ngn2 in neuronal subtype specification independently. In the transgenic models, the Ascl1 (Ngn2) coding sequence is replaced by Ngn2 (Ascl1), so the lost function of Ascl1 (Ngn2) is *rescued by Ngn2 (Ascl1)*. The mouse models are named as *Ascl1^{KINgn2}* and *Ngn2^{KIMash1}*, respectively. Importantly, the proneural factors are not ectopically

expressed as the endogenous mRNA levels are retained. In the cortex of the *Ngn2*^{KIMash1} mouse, the ventral markers are strongly up-regulated and a subpopulation of motor neurons is respecified into V2 interneurons. Interestingly, *Ngn2* fails to respecify the identity of ventral telencephalic neurons, but rescues the *Ascl1* knock-out (KO) phenotypes: normal MGE development and GABAergic neuronal specification. These experiments not only show the function of proneural genes in neural subtype specification, also indicate the difference in abilities between *Ascl1* and *Ngn2* in neuronal cell specification (Parras et al, 2002).

1.3 Epigenetic reprogramming

Embryonic stem cells (ESCs) derived from the inner cell mass (ICM) are pluripotent cells that give rise to multiple differentiated cell types during mammalian development. ESCs isolated from ICM can be cultured *in vitro* while maintaining the ability to differentiate into all cell types of the embryo proper and self-renewal activity (reviewed in Rossant, 2008). These unique advantages have made them a powerful tool to understand early embryonic development and have potential applications in regenerative medicine. Reprogramming of somatic cells is capable to generate induced pluripotent stem cells (iPSCs), which are highly similar to ESCs (Wernig et al, 2007; Chin et al, 2009). This has attracted huge amount of attention because it not only provides easy access to experimental and medical resource, also becoming a tool for understanding molecular properties of the pluripotent stem cells. More recent studies using similar approach have shown that direct conversion between somatic cells and neuronal cells can be achieved (reviewed in Yang et al, 2011; Vierbuchen and Wernig, 2012), which has important applications in studying neuronal development and regenerative medicine.

1.3.1 From somatic cells to pluripotent cells

A few approaches that promote epigenetic reprogramming of donor cells to an embryonic stem cell state have been developed over the last decade, such as somatic cell nuclear transfer (Tada et al, 2001) and fusion of somatic cells with ES cells (Cowan et al, 2005). The major breakthrough in the field is the

milestone strategy published by Takahashi and Yamanaka in 2006. By forced expression of only four TFs: Oct4 (Pou5f1), Sox2, Klf4 and Myc (c-Myc) (OSKM, also known as Yamanaka factors), somatic mammalian cells can be reprogrammed into induced pluripotent stem (iPS) cells (Takahashi and Yamanaka, 2006). Later studies are showing that some of the original Yamanaka factors are expendable and can also be replaced by other factors. For instance, Klf4 and Myc can be replaced by ESC-specific RNA binding factor Lin28 and homeobox TF Nango (Yu et al, 2007); in the presence of orphan nuclear receptor Esrrb, only Oct4 and Sox2 are required for mouse fibroblast reprogramming (Feng et al, 2009).

However, one of the major disadvantages of the iPSC generation is the low efficiency. A quantitative reprogramming experiment that plate single pre-B cells into individual culture wells shows only 3-5% of hundreds of single cell clones show pluripotency within 2 weeks and only a small subpopulation of daughter cells become fully reprogrammed (Hanna et al, 2009). In recent years, a number of studies have shown that the reprogramming efficiency can be improved by applying different factors or small molecules. The active Wnt signalling pathway is involved in the maintenance and proliferation of stem cells (Willert et al, 2003). As one of the downstream target, glycogen synthase kinase-3 (Gsk3) is inhibited by the activation of the Wnt signalling pathway. Furthermore, inhibition of mitogen activated protein kinase/Erk kinase (Mek) pathway represses lineage commitment and maintains self-renewal ability of ESCs (Ying et al, 2008). Application of Wnt3a ligands or PD0325901+CHIR99021 (2i) (inhibitor of Mek and Gsk3 pathways) significantly improves the reprogramming efficiency (Marson et al, 2008; Silva et al, 2008). Moreover, inhibition of the tumour suppressor p53 by siRNA increases the reprogramming efficiency by 100-fold (Zhao et al, 2008). These data indicate that a high cell proliferation rate is essential for the generation of iPSCs (Ruiz et al, 2011).

On the other hand, altering the epigenetic landscape also enhances the generation of iPSCs. DNA methyltransferases (Dnmts) are enzymes involved in the catalyzation and the maintenance of DNA methylation, which silence the gene transcription (reviewed in Jin et al, 2011). Dnmt inhibitors 5-azacytidine (5-AZA), RG108 and Dnmt shRNAs not only enhance reprogramming efficiency

and promote the full reprogramming of iPSCs, can also replace Sox2 or/and Myc (Mikkelsen et al, 2008; Shi et al, 2008). Histone acetylation is another critical regulation mechanism for gene transcription, which increases the accessibility of TFs to the chromatin and increase transcriptional activity (reviewed in Struhl, 1998). The application of histone deacetylase inhibitors such as valproic acid (VPA) also shows significant improvement in reprogramming efficiency (Huangfu et al, 2008a; Huangfu et al, 2008b). Taking together, it has become clear that change in epigenetic landscapes is essential for a successful reprogramming process.

1.3.2 Direct generation of neuronal cells from somatic cells

Direct conversion between two lineages was first demonstrated in the late 1970's. DNA demethylating agent 5-azacytidine is sufficient to induce myogenic, chondrogenic and doipogenic clones from an immortalized embryonic mouse fibroblast line (Taylor and Jones, 1979). Following the identification of proneural bHLH factors, early studies have demonstrated that forced expression of Ngn1 can induce pan-neuronal marker expression and immature neuronal morphologies in chick embryonic fibroblasts (Perez, et al, 1999). More recently, the Wernig group has shown that fibroblasts can be converted into functional neurons by overexpression of neuronal lineage TFs (Vierbuchen et al, 2010). Narrowing down from a pool of 19 genes, they identified a pool of three genes: Ascl1, Brn2 and Myt1l (BAM) are sufficient to induce the generation of functional neurons (named as induced-neural cells, iN cells). Surprisingly, the generation of iN is rapid (neuronal marker Tuj1 can be detected 3-5 days after induction of transgenes) and efficient (nearly 20% in 2 weeks). iN cells have the phenotypes of glutamatergic neurons, although the expression of GABAergic neuronal markers GABA and GAD67 can be detected at the early stage.

More recently, a number of groups have shown fully functional iN cells can be generated from human fibroblast cells. NeuroD1 and NeuroD2 appear to be greatly involved in this process (Pang et al, 2011; Yoo et al, 2011). Furthermore, neuronal specific micro-RNAs miR9* and miR124 seem to be able to substitute proneural gene functions (Yoo et al, 2011; Ambasudhan et al,

2011). More importantly, they are able to induce immature neuronal phenotypes at a low efficiency (Figure 1.5).

Generation of subtype specific neurons is essential for the clinical application of iN cells. By expressing additional TFs besides BAM, Son and colleges have generated Hb9-positive motor neurons from mouse embryonic fibroblasts and human ESC-derived fibroblast-like cells with the efficiency up to 10% (Son et al, 2011). More importantly, gene expression profiling analysis shows the induced motor neurons (iMNs) are similar to ESCs-derived motor neurons. The generation of midbrain dopaminergic (DA) neurons is demonstrated by two groups using different combination of TFs (Pfisterer et al, 2011; Caiazzo et al, 2011; Figure 1.5). Pfisterer and associates have generated iN cells expressing tyrosine hydroxylase (TH), aromatic L-amino acid decarboxylase (AADC) and midbrain marker Nurr1. However, these iN are not able to release dopamine. iN cells generated by Caiazzo and colleges using Ascl1, Nurr1 and Lmx1a are able to release dopamine. However, iN cells generated from human fibroblasts using the same reprogramming recipe do not express midbrain specific markers and show immature morphologies. These suggest dopaminergic neurons generated from reprogramming are not midbrain specific, which is further supported by gene expressing profiling analysis (Figure 1.5).

However, the molecular mechanisms of the direct conversion of two distant lineages remain largely unknown. Some studies suggest that TFs act upstream of chromatin remodelling pathways and activate target genes repressed by repressive chromatin modifications, such as DNA methylation and histone deacetylation (reviewed in Suzuki et al, 2008; Zaret et al, 2008; reviewed in Li et al, 2007).

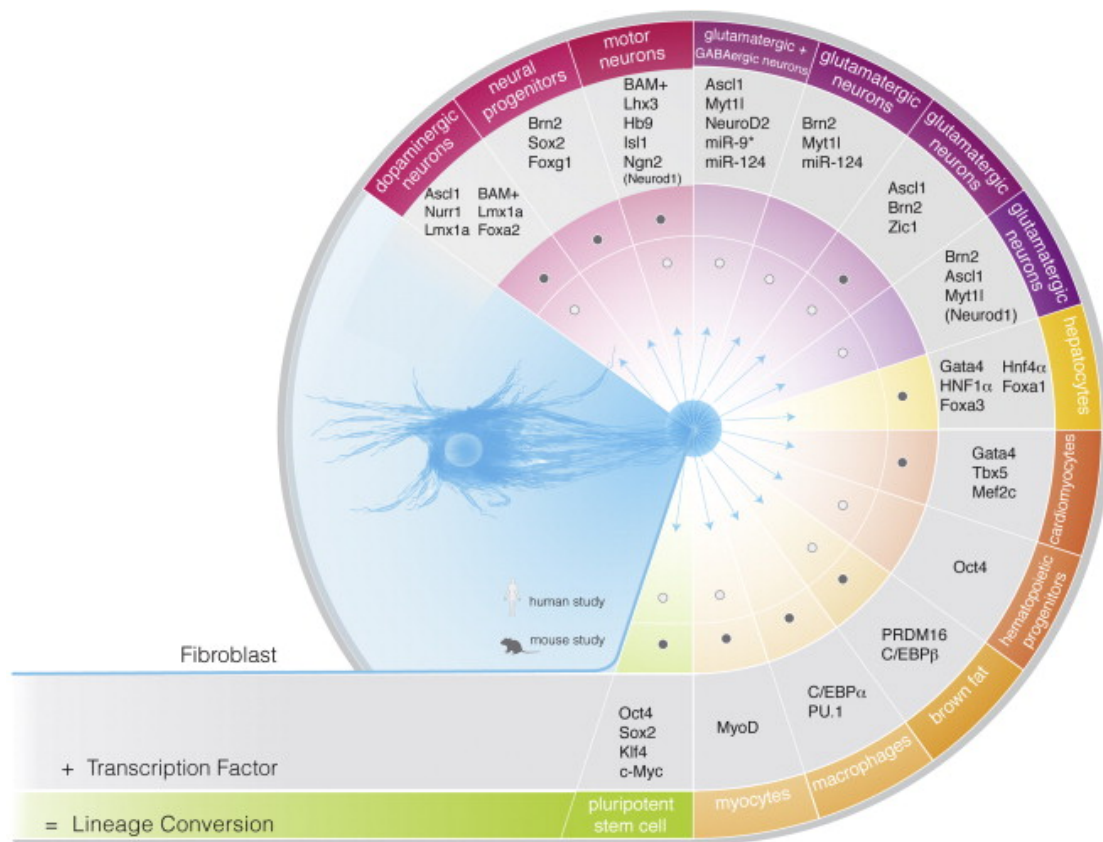


Figure 1.5: Generation of different cell lineage from fibroblasts using different combination of transcription factors
(adapted from Vierbuchen and Wernig, 2012)

1.4 Sox family TFs in cell fate specification

Sox family members are TFs contain a highly conserved high-mobility-group (HMG) domain that was first identified from the mammalian testis-determining factor Sry (Gubbay et al 1990; Sinclair et al, 1990). Generally, proteins contain a HMG domain with $\geq 50\%$ identical in amino acids to the HMG domain in Sry are defined as Sry-related HMG box (Sox) proteins. 20 Sox genes have been identified and they can be subdivided into 8 groups (namely SoxA to SoxH) based on the phylogenetic analysis of the HMG domains (reviewed in Sarkar and Hochedlinger, 2013). Sox proteins within the same group share the amino acid sequence identify at a minimum of 70%. As TFs, Sox proteins directly bind to the DNA in a sequence specific manner (core binding sequence (A/T)(A/T)CAA(A/T)G) (reviewed in Kamachi et al, 2000; reviewed in Wilson and Koopman, 2002), but with different optimal binding sequence, expertly. For example, the optimal Sox9 binding sequence is AGAACAATGG (Mertin et al, 1999).

Over the past years, functions of the Sox family TFs have been under intensive studies (reviewed in Sarkar and Hochedlinger, 2013). A number of subgroups of the Sox family have been shown to have critical functions in developing CNS. SoxB1 (Sox1-3) TFs are expressed by most of the progenitor cells in CNS and downregulated upon neuronal differentiation (Uwanogho et al, 1995; Pevny et al, 1998). Earlier studies have shown that SoxB1 TFs maintain NSC properties via counteracting activities of proneural bHLH TFs such as Ngn2 (Bylund et al, 2003). On the other hand, Sox21 (a member of the SoxB2 subgroup) co-express with Ngn2 and contracting SoxB1 activities, which subsequently promote neurogenesis (Sandberg et al, 2005). Furthermore, the SoxC subgroup is essential for establishing neuronal properties. Instead of inducing neurogenesis directly, Sox4 and Sox11 (members of the SoxC subgroup) act downstream of proneural bHLH TFs and function as transcriptional activators of neuronal genes such as Tuj1 (Bergslund et al, 2006).

The SoxE group contains three members: Sox8, Sox9 and Sox10. It has been shown that both Sox9 and Sox10 are greatly involved in stem cell maintenance and specification in vertebrate nerve system. Sox9 is involved in the formation of neural crest (Cheung and Briscoe, 2003). Recent GoF and LoF

studies also indicate the Sox9 function in specification and maintenance of NSCs by acting downstream of the Shh signalling pathway in developing mouse telencephalon (Scott et al, 2010). Furthermore, Sox9 is suggested to switch from neurogenesis to gliogenesis in developing spinal cord (Stolt et al, 2003). Sox10 is expressed in neural crest stem cells (NCSCs) and is down-regulated in NCSC derived smooth muscle cells and neurons, but remained in the NCSC derived Schwann cells (reviewed in Anderson 1997). GoF studies suggest Sox10 is essential for NCSCs maintenance by preventing BMP2-induced neural and smooth muscle differentiation and TGF β -induced proliferative arrest (Kim et al, 2003). On the other hand, Sox10 can interact with Oct6 to induce the expression of Krox20 (Ghislain and Charnay, 2006). The synergised expression of Sox10 and Krox20 activates myelin specific genes and lead to terminal differentiation and myelin formation of Schwann cells (Bondurand et al, 2001; LeBlanc et al, 2007). Comparing with the other two members of the group, the function of Sox8 in developing nerve system is much less clear. Two studies indicate that Sox8 functions redundantly with Sox9 and Sox10 in neural crest development in *Xenopus* (O'Donnell et al, 2006) and is involved in the oligodendrocyte development in a Sox10 dependent manner in the spinal cord (Stolt et al, 2004). In other tissues, Sox8 is shown to prevent MyoD-mediated myogenesis (Schmidt, et al, 2003) and negatively regulated osteoblast differentiation during bone formation (Schmidt et al, 2005).

1.5 Cell cycle progression and regulation

Cell cycle and its phases were first described by Howard and Pele in 1951 and it became clear this is a highly conserved and regulated process. A complete cell cycle can be divided into interphase (G1-, S- and G2-phase) and mitotic phase (M-phase). Interphase is the stage that a cell spends most of its time during cell cycle to duplicate cellular contents (G1-phase) and chromosomes (S-phase) and prepare for mitosis (G2-phase). The mitotic phase (M-phase) is a short stage comparing with the interphase and to produce two identical daughter cells through cytokinesis (Figure 1.6 A). During mitosis, spindles attach to duplicated chromosomes (prophase), which are aligned around metaphase plate (prometaphase and metaphase). The centromeres then divide and sister

chromatids are pulled by spindles towards the opposite poles of the cells (anaphase). The nuclear member reforms around the decondensed chromosomes (telophase) and two daughter cells are formed after cytokinesis (Figure 1.6 B).

Cell cycle is strictly controlled by a number of regulatory mechanisms that either promote or inhibit cell proliferation. Cyclin dependent kinases (Ckds), cyclins and Cdk inhibitors (CkdI) are the families provide primary regulatory machinery to control cell cycle progression.

The Cdk family members contain a serine/threonine catalytic domain and associate with regulatory partner cyclin, a protein contains the cyclin box that mediate binding to Cdk (Gopinathan et al, 2011), to promote its kinase activity. The Cdk/cyclin complex drives cell cycle progression and promotes phase transition. For example, Cdk1/Cyclin B1 complex phosphorylate FoxM1 TF, which is essential for chromosomal segregation and cytokinesis, and subsequently drive M-phase progression (Costa, 2005; reviewed in Chen et al, 2009). On the other hand, CdkI binds and prevents phosphorylation of Cdk to block cell cycle progression. Based on the structure and Cdk binding specificity, CdkI is divided into two families: the Ink4 family members p16^{INK4a}, p15^{INK4b}, p18^{INK4c} and p19^{INK4d} primarily target Cdk4 and Cdk6. The Cip/Kip family members p21^{Cip1}, p27^{Kip1} and p57^{Kip2} are more broadly interfere with the activities of cyclin D-, E-, A- and B-dependent kinase complexes (Sherr and Roberts, 1999).

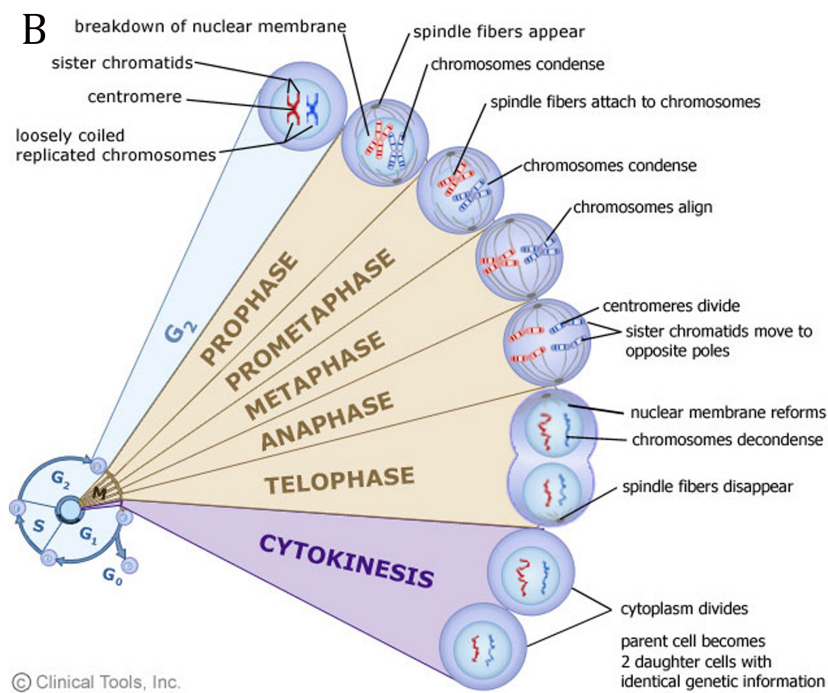
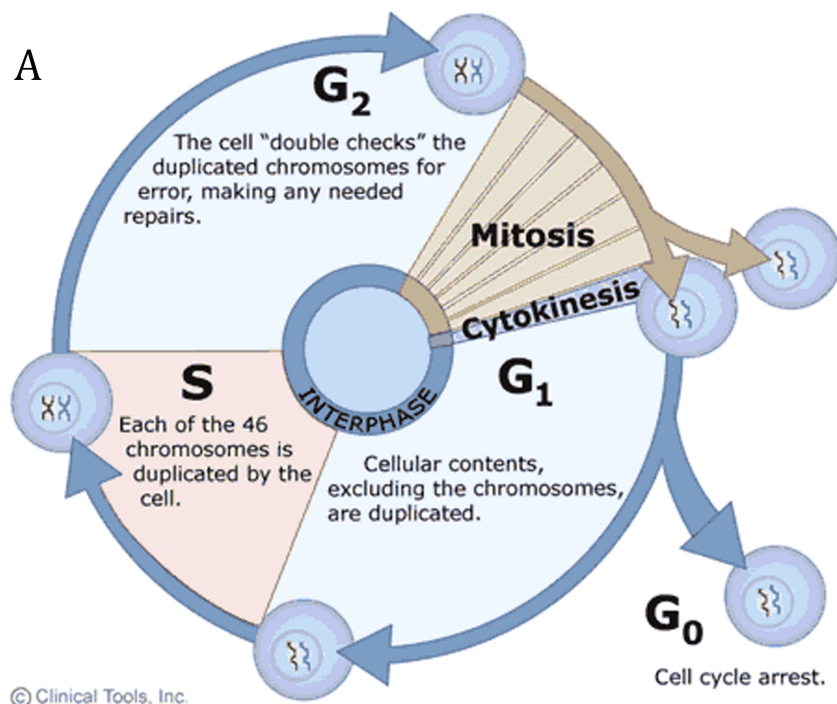


Figure 1.6: Different stages of cell cycle (A) and mitosis (B)
(source: Clinical Tools inc)

Chapter 2: Material and method

2.1 Cloning

All plasmids were prepared and analysed using standard maxiprep (NucleoBond PC500, Macherey-Nagel, 740574.25) and miniprep (GelElute Plasmid miniprep Kit, Sigma, PLN350-1KT) kits according to manufacturer's instruction. Restriction digestions were carried out according to Current Protocols in Molecular Biology using Buffer and Restriction Enzymes from New England Biolabs or Roche. Restriction digests were purified using QIAquick Gel extraction Kit (Qiagen, 28706) according to manufacturer's instruction. Plasmids were grown in OneShot Top10 chemically competent *Escherichia Coli* cells in Luria Bertani (LB) medium or LB/Agar plates (for single colony isolation) containing 100mg/ml ampicillin (Sigma).

2.2 Generation of short hairpin (sh)RNA containing plasmids

Oligonucleotides for shRNA were designed by online tool Whitehead RNAi (<http://sirna.wi.mit.edu>) or ordered based on published data (see 2.9.1 for sequences of oligonucleotides).

Oligonucleotides were ordered from Sigma and dissolved in water with the concentration of 100 μ M (100pmol/ μ l). Each hybridization reaction contained 800pmol of forward and reverse shRNA oligonucleotides, 2 μ l of 10X NEB3 buffer and 2 μ l of water. Sample was heated at 95C° for 5 minutes and allowed to cool down to room temperature in the heat block.

P-Super-Puro and P-Super-neo-GFP vectors were digested with BglII and XhoI restriction enzymes and purified by gel extraction kit according to manufacturer's instruction. Hybridized shRNA was ligated with the digested vector using T4 DNA ligase and buffer (P-Super-Puro-shRNA and P-Super-neo-GFP-shRNA). The correction insertion was confirmed by sequencing (Sourcebioscience).

P-Super-Puro-shRNA and p-Ad5-GFP lentiviral vector were digested with XhoI and EcoRI restriction enzymes. The 350bp fragment from P-Super-Puro-shRNA that contained the shRNA was ligated with digested p-Ad5-GFP lentiviral

vector using T4 DNA ligase and buffer (p-Ad5-shRNA-GFP). The correction insertion was confirmed by sequencing.

2.3 Production of lentivirus

Lentiviruses were produced from HEK293T cells. Lipofectamine 2000 (Invitrogen, 11668-019) was used for cell transfection.

3 million HEK293T cells were seeded in 10cm cell culture dish the night before the transfection. For each transfection reaction, 12µg of lentiviral plasmids, 9µg of psPax2 lentiviral packaging plasmids and 3µg of VSV-G pseudotyping plasmids were mixed with 1.5ml opti-MEM medium; 60µl of Lipofectamine 2000 reagent was gently mixed with 1.5ml opti-MEM medium and incubated at room temperature for 5 minutes. The plasmids mixture and Lipofectamine 2000 mixture were gently mixed and incubated at room temperature for 20 minutes; HEK293T cells were carefully washed with PBS before the transfection reagent was applied and incubated at 37C° for 4.5 hours; the transfection reagent was then replaced with pre-warmed HEK-293T medium.

The supernatant was collected at 24h and 48h. Lentiviral particles were concentrated from the supernatant by ultra-centrifugation at 20,000rpm for 2 hours. The viral pellet was resuspended in PBS and stored at -80C°.

2.4 RNA isolation and cDNA synthesis

Cells for total RNA isolation were prepared in 6-well plate. Growth medium was removed from the well and 1ml TRIzol reagent (Invitrogen, 15596-026) was added directly to the cells. The cells were lysed by pipetting up and down several times and collected in the 1.5ml eppendorf tube. The homogenized sample was incubated at room temperature for 5 minutes. 200µl of chloroform was added and the sample was shaken vigorously for 15 seconds before incubated at room temperature for 2 minutes. Phases were separated by centrifugation at 12,000g for 15 minutes at 4C°. The aqueous upper phase was carefully transferred into a new tube and 500µl of 100% isopropanol was added for RNA precipitation. The sample was mixed and incubated at room temperature for 10 minutes. The sample was then centrifuged at 12,000g for 10

minutes at 4C° and the RNA pellet was washed with 1ml 75% ethanol and subsequently centrifuged at 12,000g for 5 minutes at 4C°. The pellet was air-dried and dissolved in 50µl of RNase free water. The concentration of the RNA was checked by a spectrophotometer. For expression microarray analysis, RNA cleanup was performed using RNeasy Mini Kit (QIAGEN, 76104) according to the manufacturer's instruction.

cDNA was synthesised using the High Capacity RNA-to-cDNA master mix (Applied Biosystem, 4390781). For each 20µl of reverse transcriptase-PCR (RT-PCR) reaction, 400ng of total RNA and 4µl of RNA-to-cDNA master mix were used. The reaction was incubated at 25C° for 5 minutes and then 42C° for 30 minutes. The reaction was stopped by heating to 85C° for 5 minutes and hold at 4C°. The cDNA product was diluted 12 times for further use.

2.5 Quantitative real time PCR

Quantitative Real time PCR (q-PCR) was performed according to the manufacturer's instruction using iTaq SYBR Green supermix with Rox (Bio-Rad, 172-5853) on an ABI7900 machine (Applied Biosystem). For each 16µl reaction, 8µl of SYBR Green supermix, 3µl of water, 1µl of primer mix (180µl water mixed with 10µl of forward and 10µl reverse primers, (see 2.9.1 for primer sequences) and 4µl of cDNA template (see 2.4 for cDNA synthesis). The housekeeping genes Tbp and Hprt were used for normalization. The relative expression was calculated as $E = 2^{-\Delta Ct}$, where $\Delta Ct = Ct_{\text{Gene Analysed}} - Ct_{\text{Housekeeping Gene}}$ (Ct: mean threshold of cycle number).

2.6 Expression microarray analysis

For microarray analysis, the concentration and quality of the total RNA was assessed by spectrophotometry (RNA sample requirement: minimum 60ng of RNA after RNA cleanup; the 260/230 ratio ≥ 1.8). All RNA samples were diluted to 30ng/µl and processed according to the Illumina Whole-Genome Gene Expression Direct Hybridization Assay Guide using Illumina® TotalPrep™-96 RNA Amplification Kit. Samples were hybridized to a Mouse Ref-8 v2 arrays and scanned by the BeadArray Reader. The array intensity data was loaded into the Illumina BeadStudio software and normalized with the background. All analysis

was performed on biological triplicates. The significant analysis was performed by GeneSpring software and using one-way-ANOVA with subsequent p-value correction by Benjamini-Hochberg post-hoc correction.

For MEFs Ascl1 GoF array, significantly de-regulated probes were defined as probes with a corrected p-value ≤ 0.05 and a fold change ≥ 2 related to eGFP viral infected cells at same time point. For NS5 SoxE LoF array, significantly de-regulated probes were defined as probes with a corrected p-value ≤ 0.05 and a fold change ≥ 1.5 related Con shRNA transfected cells at the same time point.

2.7 *in vitro* cell culture

2.7.1 P19 and human embryonic kidney (HEK) 293T cell culture

P19 embryonic carcinoma cells HEK293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Sigma, D5671) supplemented with 10% fetal calf serum (FBS, Hyclone), 50IU Penicilin/Streptomycin (Invitrogen) and 2mM L-Glutamine (Invitrogen).

2.7.2 NS5 cell culture

NS5 cells were cultured in EUROMED-N medium (EURO-LONE, ECM0883L) supplemented with 50IU Penicilin/Streptomycin (Invitrogen), 2mM L-Glutamine (Invitrogen), N2-supplement (R&D, AR009), Bovine Serum Albumin (BSA), EGF (peprotech, 315-09, 10ng/ml), FGF-2 (peprotech, 450-33, 10ng/ml) and Laminin (sigma, L2020, 20ug/ml).

2.7.3 Mouse embryonic fibroblasts (MEFs) isolation and culture

MEFs were isolated from E14-E14.5 wild-type Parkes mice embryos. The head, vertebral column, dorsal root ganglia and all internal organs were carefully removed and discarded. The remaining tissues were dissociated with 0.25% trypsin (Sigma) at 37C° for 10-15 minutes to create single cell suspension. The cells were cultured in T150 tissue culture flask in MEF medium [DMEM supplemented with 10% FES, non-essential amino acid (Invitrogen), sodium pyruvate (Invitrogen) and 50IU Penicilin/Streptomycin] at 37C° until confluent. The cells were frozen and stored in liquid nitrogen.

2.7.4 Direct reprogramming of MEF to neurons

Lentiviral plasmids Tet-O-FUW-Ascl1, Tet-O-FUW-Brn2 and Tet-O-FUW-Myt1l were kindly provided by Dr Vania Broccoli, San Raffaele Institute, Milan, Italy. Tet-O-FUW-eGFP plasmid was purchased from addgene (addgene plasmid 30130). Tet-O-FUW empty plasmid was generated from Tet-O-FUW-eGFP by EcoRI restriction enzyme digestion. The expression of the transgenes was doxycycline inducible (Figure 3.1 B-E).

Protocol for direct reprogramming of MEF into neurons by transcriptional factors (Ascl1, Brn2 and Myt1l) was established by Marius Wernig group (Vierbuchen et al, 2010). Briefly, Ascl1, Brn2 and Myt1l transgenes were delivered to cultured MEFs (see 2.7.2 for MEFs isolation) by lentiviral infection (see 2.3 for production of lentiviruses). The expression of the reprogramming factors was induced by the application of doxycycline. Ascl1, Brn2 and Myt1l lentiviruses or Ascl1 lentiviruses were added to passage 3 MEFs, which were seeded the night before the infection (60,000/well for 24-well plate). The total amount of viruses added into each condition remained the same by adding viruses carrying empty lentiviral vector. After 20-24 hours in medium containing lentivirus, the cells were switch to fresh MEF medium containing doxycycline (2µg/ml). 48 hours later, the culture medium was switched to N3 medium [DMEM/F12 (Invitrogen) containing insulin (25µg/ml, Sigma), transferrin (50µg/ml, Sigma), sodium selenite (30nM, Sigma), progesterone (20nM, Sigma), putrescine (100nM, Sigma) and Penicilin/Streptomycin (50IU, Simga)] containing doxycycline (Sigma). The N3 medium was changed every 2-3 days. Reprogrammed MEFs were fixed at D-7, D-13 and D-22 for analysis.

2.8 Immunocytochemistry and image analysis

For *in vitro* immunostaining, cells were cultured on glass coverslips (Fisher Scientific). Cells were fixed with 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS) containing 15% sucrose for 25 minutes at room temperature. The cells were then permeabilized with 0.5% Triton (Sigma) for 8 minutes and blocked with 10% FBS. Incubation with primary and secondary antibody was performed at room temperature for 1.5 hours and 45 minutes (see 2.9.2 for antibody used in experiments)

Immunofluorescence images were captured using a ProgRes C14 camera (Jenoptik) linked to an Axioplan II microscope (Zeiss) under the 20X objective. For MEF reprogramming experiments, 10 images were randomly taken for individual culture conditions. Images were processed and analysed with Openlab (Perkin Elmer) and Fiji (NIH).

2.9 Western blot

Immunoblots were done following the standard protocol. Equal amount of protein were separated on a 10% SDS-PAGE and blotted into nitrocellulose membranes (BioRad). Antibodies against Sox9, V5 and actin (loading control) were used to detection (see 234 for antibody dilution). For secondary antibodies, horseradish peroxidase (HRP)-linked anti-mouse/rabbit antibodies were used at a 1:2000 dilution.

2.10 Fluorescence-activated cell sorting (FACS) analysis

Cultured cells were collected and resuspended in ice cold 70% EtOH. Samples were immediately put on ice then transferred to -20 C° for 2h.

Cells were resuspended in 0.5ml PBS supplemented with RNaseA (0.1mg/ml) and Propidium iodide (sigma, 40ng/ml). Samples were incubated at 37 C° for 1h and kept on ice until analyse.

2.11 Materials

2.11.1 Oligonucleotides

2.11.1.1. Cloning

Aof1 shRNA1 FW:

TCGAGAAAAAGGCGAGGCCTATGATATTATCTCTTGAATAATATCATAGGCCTCGCC
GGG

Aof1 shRNA2 FW:

TCGAGAAAAAGGTACTGATAAGGGAAGTATCTCTTGAATACTTCCCTTATCAGTACC
GGG

Cbfa2t3 shRNA1 FW:

GATCCCCGCGTCTCTGTCTTTGTAAATTCAAGAGATTACAAAGACAGAGACGCTTT
TTTC

Cbfa2t3 shRNA2 FW:

GATCCCCGTGGCGATCAGGTTAATTATTCAAGAGATAATTAACCTGATCGCCACTTT
TTC

Sox8 shRNA FW:

GATCCCCGCAGCGAGGTTATCAGTAATTCAAGAGATTACTGATAACCTCGCTGCTTT
TTC

Sox9 shRNA FW:

GATCCCCGTGTGTTGTTATAGTAACATTCAAGAGATGTTACTATAACAACACACTTT
TTC

Zfp238 shRNA1 FW:

GATCCCCGCTTCATGCAGCATGTATTTCTTCAAGAGAGAAATACATGCTGCATGAAG
CTTTTTC

Zfp238 shRNA2 FW:

GATCCCCGGGTATCAGACGTGTTACATTCAAGAGATGTAACACGTCTGATACCCTTT
TTC

2.11.1.2 gene expression analysis

Aof1: AGATGGAGCCCTCCAAGG(FW) GCACTTCTGAGGGGTGAGAG(RV)

Ascl1 endo: TCTCCTGGGAATGGACTTTG(FW) GGTGGCTGTCTGGTTTGT(RV)

Ascl1 ORF: CGTCCTCTCCGGAAGTATG(FW) CTGCTGTGGCAGGCTGTAG(RV)

Cbfa2t3: ACCCTCCTTCACACCTCACA(FW) GCTGAAGCCATTGGGTGTA(RV)

Dpysl4: GACCCAGCTGATGATTTCT(FW) ACCAGCGTCAGGAAACACAT(RV)

Fjx1: GCTTGTGAGCAACCTCTTCA(FW) TCGGAATACACACACCGACT(RV)

Hapb4: CCTGGGGATCAGGTAAAGAC (FW) GTGCAGGTGGCTCTGTATCA (RV)

Hes6: ACCCACCTCGCTGAGTAGTT(FW) CCTGAAGACTCTCGTTGATCC(RV)

Hprt: AGCCTAAGATGAGCGCAAGT (FW) ATGGCCACAGGACTAGAACA(RV)

Id2: CAGAGACCTGGACAGAACCA (FW) TCAGAAGGGAATTCAGATGC (RV)

Jmjd2c: CAGCAGGTAGCGAGTGATGA(FW) TGCTGTCTCCCATCTAGAATCA(RV)

Mfng: TCGTGGTCACCAACTGTTCT(FW) GCCTTGGGGTTCACATAGTT(RV)

Nav1: TCTACAGGAAGGAGCCAAGG(FW) GTCTTCCAAGCAGCTTTCT(RV)

Ncoa1: GGAACCTTTTCACCCAACAT (FW) TGCAAACCTGGTTCAAGATGG (RV)
 Nes: GAGGTGACCCTTGGGTTAGA (FW) TAGCCCTACCACTTCCTGCT(RV)
 Plekhg5: GAGCTACCCTCCCCAGACTC(FW) TGTGGAACCTTGCTATCACAGG(RV)
 Sema6b: GACACTGTTCATTGGGGACA(FW) GATGTCACTGGGGTTGGAG(RV)
 Sox8: GTACCCGCATCTCCATAACG (FW) GGCCCAGTTCAGTACCAGAG(RV)
 Sox9: AGAACAAGCCACACGTCAAG(FW) TCCACGAAGGGTCTCTTCTC(RV)
 Sox10: CAGCAAGACACTAGGCAAGC(FW) CTTGTAGTCCGGATGGTCCT(RV)
 Tbp: GGGGAGCTGTGATGTGAAGT(FW) CCAGGAAATAATTCTGGCTCA(RV)
 Zfhx1a: AAGAACTGCTGGCAAGACAA(FW) GCTGCAGAAATTCTTCCACA(RV)
 Zfp238: GATGGTCCCCAGTGAGAGAA(FW) GACAGTCCCAGTGACCCAGT(RV)

2.11.2 Antibody and dilution

2.11.2.1 Immunocytochemistry

Mouse anti-Neuronal Class III β -Tubulin (Tuj1): Vovance (MMS-435P): 1:400
 Rat anti-Microtubule Associated Protine 2 (MAP2): Millipore (AB5622): 1:400
 Rabbit Anti-enhanced Green Fluorescent Protein (eGFP): innitrogen (A11122):
 1:500
 Rabbit anti-Glial Fibrillary Acidic Protein (GFAP): Dako (Z0334): 1:1000
 Goat anti-Doublecortin (DCX): Santa Crus (sc-8066): 1:50
 Rabbit anti-NG2 Chondroitin Sulfate Proteoglycan: Millipore (AB5320): 1:500
 Rabbit anti-Phospho Histone H3 (pHH3): Millipore (06-570): 1:500

2.11.2.2 Western blot

Rabbit anti-Actin: sigma (2066):1:1000
 Mouse anti-V5: Invitrogen (R960-25): 1:2000
 Rabbit anti-Sox9: Millipore (AB:5535): 1:400

Chapter 3: Functions of Sox8 in direct conversion of mouse embryonic fibroblasts to neurons

3.1 Rational, hypothesis and aim

Ascl1 (also known as Mash1) is a basic-Helix-Loop-Helix (bHLH) transcriptional factor (TF), which has critical functions in neuronal differentiation and cell type specification during central nerve system (CNS) development. Ascl1 is primarily expressed at the ventral region of developing telencephalon in mouse (Fode et al, 2000), which is the birthplace of GABAergic neurons (Parras et al, 2006). Here, neural stem cells expressing Ascl1 are withdrawn from cell cycle and committed to neuronal differentiation. Genome-wide studies using ChIP-on-Chip arrays have identified Ascl1-binding sites in the proximal promoter regions, which suggest that Ascl1 regulates genes involved in all major steps of neurogenesis in a direct manner (Castro et al, 2010).

In 2010, Wernig and associates established the protocol to directly convert fibroblasts into functional neurons by overexpression of neurogenic TFs: Ascl1, Brn2 and Myt1l (Vierbuchen et al, 2010). Ascl1 is the essential factor to ensure the efficient reprogramming process. Using the similar approach, a number of groups using different TF recipes have reprogrammed mouse or/and human fibroblasts into functional neurons (reviewed in Vierbuchen and Wernig, 2012). Ascl1 is the critical factor involved in the majority of the recipes. Furthermore, ectopic expression of Ascl1 alone is sufficient to induce the generation of immature neurons (Figure 3.1 D and 1D'; Vierbuchen et al, 2010). These suggest that Ascl1 has critical functions in the direct reprogramming mechanisms.

As Ascl1 is an essential factor to activate neuronal specific programme in both neuronal differentiation and direct reprogramming, we hypothesise that identifying common Ascl1 downstream genes allows us to understand molecular mechanisms in both programmes.

Here, I used Ascl1 Gain of Function (GoF) approach to identify Ascl1 downstream genes at the early stage of MEFs reprogramming and further

investigated functions of Sox8 in Ascl1-mediated reprogramming process using both GoF and Loss of Function (LoF) studies.

3.2 Re-establishment of MEFs reprogramming *in vitro*

In order to identify downstream factors contributing to Ascl1-mediated reprogramming, it was essential to demonstrate the *in vitro* reprogramming system could be re-established. To reproduce the direct reprogramming experiment, I followed the protocol established by the Wernig group (Vierbuchen et al, 2010). Briefly, lentiviruses carrying Ascl1, Brn2 and Myt1l transgene were used to infect passage 3 MEFs isolated from E14-E14.5 wild-type Parkes mouse embryos. The expression of the transgene was doxycycline dependent. Induced neurons (iNs) were identified by immunocytochemistry using antibody against neuronal class III β -tubulin (Tuj1) (Figure 3.1 A and B). In addition to the three-TF condition (Ascl1+Brn2+Myt1l; BAM condition), I also used Ascl1 alone to reprogram MEFs for two reasons. Firstly, previous studies have shown that Ascl1 is sufficient to induce immature neuronal features (Vierbuchen et al, 2010), which suggests that Ascl1 is sufficient to activate neuronal specific program during direct reprogramming of fibroblasts. Secondly, I was focusing on the functions of Ascl1 during the reprogramming process and the additional expression of Brn2 and Myt1l might interfere the identification of Ascl1 downstream genes.

At day 7, Tuj1-positive cells were detected in both BAM (Figure 3.1 C) and Ascl1 alone (Figure 3.1 D) conditions. To compare the morphology of the reprogrammed cells, it was clear that cells reprogrammed under the BAM condition had a more mature neuronal phenotype (cells had bipolar or multipolar structure and long extended processes; Figure 3.1 C and D). The observation was also confirmed by the quantification. Tuj1-positive cells were divided into 2 populations: one with processes longer than 3 cell body length (mature neurons, which was defined as iN) and the other one with processes shorter than 3 cell body length (immature neurons). 70% of the Tuj1-positive cells could be defined as iN under the BAM condition and only 12.5% of Tuj1-positive cells under Ascl1 only condition could be classified as iN (Figure 3.1 E). The reprogramming efficiency (calculated using only the number of iN) was 15%

under the BAM condition and 5% using Ascl1 alone (Figure 3.1 E'). Similar results were observed at day 13 cultures (Figure 3.1 C' and D'). Under the BAM and Ascl1 only conditions, 90% and 15% of Tuj1-positive cells could be defined as iN and the reprogramming efficiency was 12% and 6% (Figure 3.1 F and F').

In conclusion, we have successfully re-established the *in vitro* reprogramming system by converting MEFs into neurons.

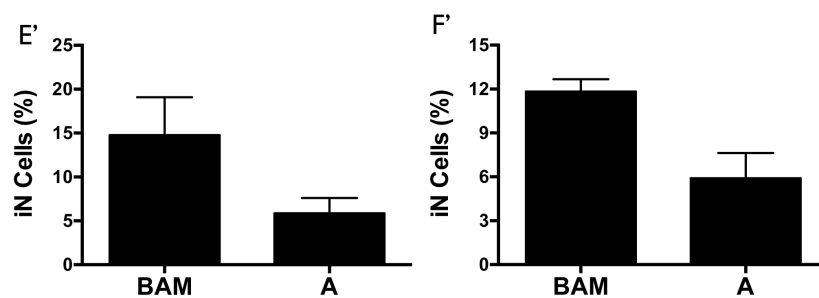
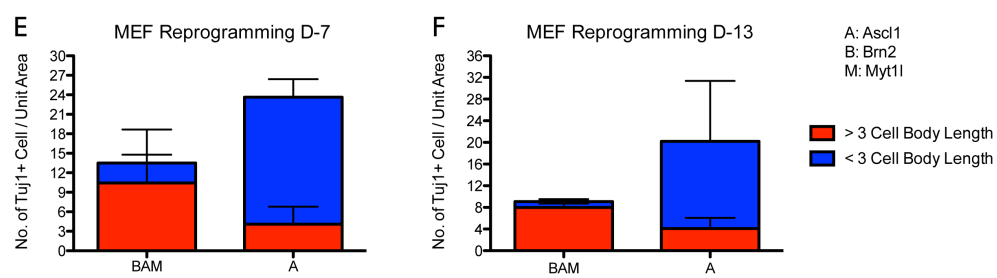
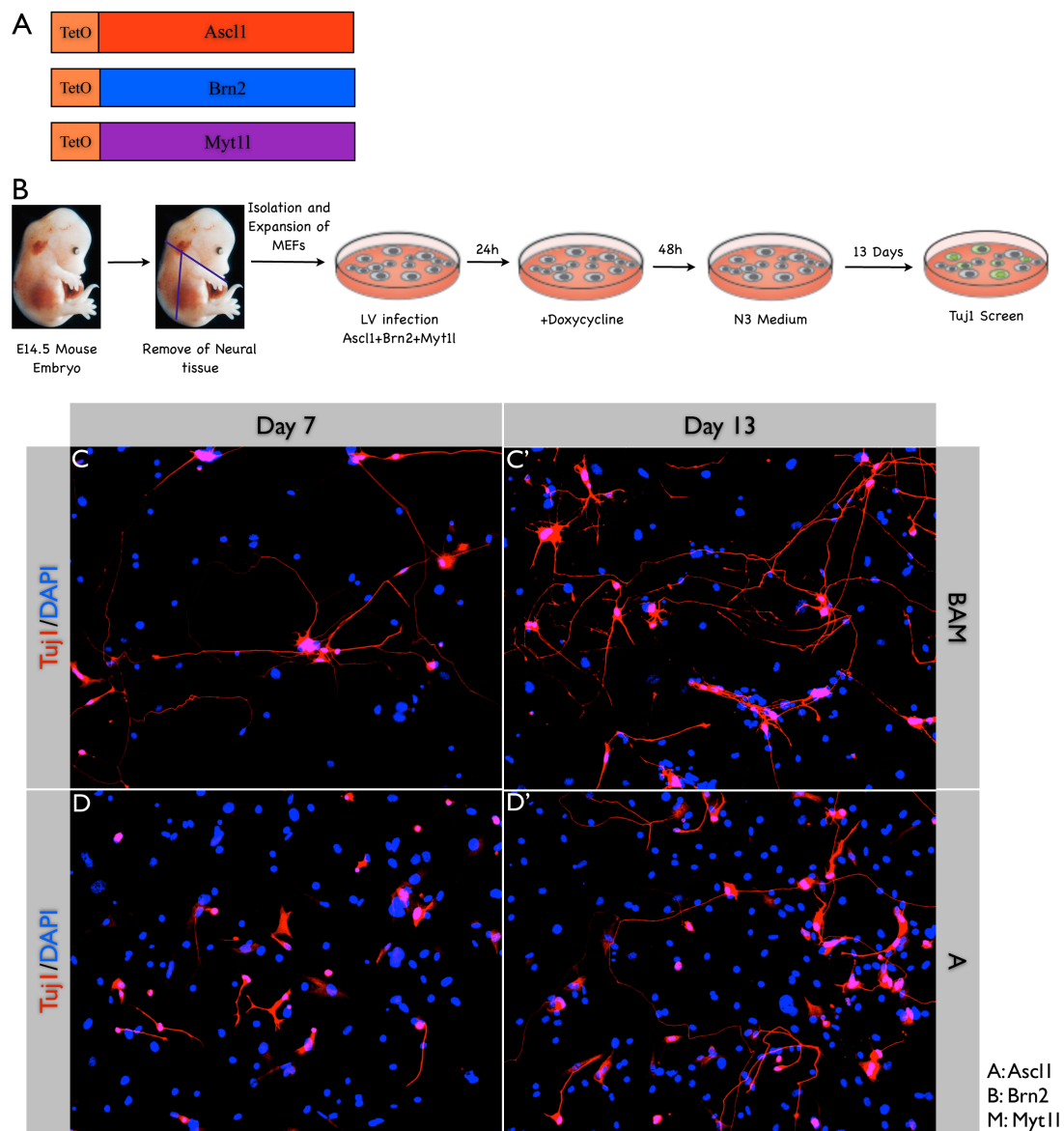


Figure 3.1: Direct conversion of MEFs into neurons

Lentiviral plasmids carrying *Ascl1*, *Brn2* and *Myt1l* transgenes. The expression of the transgene is doxycycline dependent (A). Experimental rationale (B). P3 MEFs were infected with viruses carrying three-TFs (BAM) or *Ascl1* alone. MEFs iN express Tuj1 (red) and display complex neuronal morphologies 7 and 13 days after infection under the BAM condition (C and C'). *Ascl1* alone induces Tuj1 expression, but the majority of the Tuj1-positive cells display immature neuronal morphologies (D and D'). The quantification graphs show the number of Tuj1-positive cells/unit area under 20x objective and the reprogramming efficiency (calculated using only iNs) under BAM or *Ascl1* alone conditions 7 days (E and E') and 13 days (F and F') after infection.

3.3 Identification of Ascl1 downstream genes at the early stage of MEF reprogramming

To assess whether common regulatory mechanisms were involved in both Ascl1-mediated MEFs reprogramming and neuronal differentiation, I decided to perform genome-wide expression microarrays to detect gene expression changes at the early stage of MEFs reprogramming under GoF condition. As shown in earlier experiments, Ascl1 alone is sufficient to induce the activation of neuronal specific programs (Figure 3.1 D and E). Furthermore, I aimed to identify Ascl1 downstream genes and additional expression of Brn2 and Myt1l would interfere the experiment. For these reasons, the gene expression microarrays were performed under Ascl1 GoF condition. Two experiments were performed for mRNA sample collection (Figure 3.2 A). Experiment A: lentiviruses carrying Ascl1 transgene and doxycycline were added into P3 MEFs cultures at the same time. RNA samples were collected at 12h (A1) and 24h (A2). Experiment B: doxycycline was added into the cell culture 24h post infection. RNA samples were collected at 12h (B1) and 24h (B2). Samples were collected from three independent biological replicates.

Delta-like 3 (Dll3) has been shown to be a direct target of Ascl1 and was used as a read-out for Ascl1 function in our *in vitro* system (Castro et al, 2006). The expression of Dll3 and Ascl1 was analysed by quantitative RT-PCR (qPCR) as quality control of the mRNA samples. Compared to eGFP infected control samples, significant up regulation of Ascl1 expression was observed (A1, A2, B1 and B2). Dll3 expression was up regulated in A2, B1 and B2. Furthermore, the expression level of Ascl1 and Dll3 were had minor differences between biological replicates (Figure 3.2 B). Nevertheless, addition of doxycycline led to an increase in Ascl1 mRNA expression and up regulation of early Ascl1 target Dll3 suggested Ascl1 GoF was sufficient to deregulate downstream genes in our *in vitro* system.

To assess and follow gene expression changes during Ascl1-induced MEFs reprogramming globally, I took advantage of transcriptome-wide expression microarrays. The experimental setup and quality control were described previously. mRNA samples were hybridized to Illumina Ref8v2 BeadArrays (performed by Harsha Jani, NIMR sequencing facility) and analysed using the GeneSpring software. The minimum requirements for significantly deregulated

(SD) genes were a corrected p-value lower than 0.05 and a relative fold change higher than 2. As shown in Figure 3.2 A, the number of SD genes was determined for each experiment. In experiment A1, only 3 genes fulfil the minimum requirements, the data was not considered for further analysis. 246 genes were commonly deregulated in A2, B1 and B2 (205 up regulated and 41 down regulated, Figure 3.2 C). In order to validate the quality of the gene expression profiling data, 5 SD genes were cherry picked from the pool of 205 up regulated genes and the expression level were validated using qPCR. Samples used for expression microarrays were used for qPCR validation (the expression of Atoh8 and Sox8 were tested for experiment B1 and B2 due to lack of mRNA materials). All genes validated showed significant up regulation under the Ascl1 GoF condition, which suggested the gene expression profiles were reliable (Figure 3.2 D).

To have global understanding of functions of Ascl1 downstream genes at early stage of MEFs reprogramming, the 246 commonly SD genes were annotated to functional clusters using Gene Ontology (GO) analysis (<http://david.abcc.ncifcrf.gov/>). The top ten functional clusters for up regulated (Figure 3.3 A) and down regulated genes (Figure 3.3 B) were manually grouped based on boarder GO terms. No functional cluster related with neuronal development was observed from GO analysis (Figure 3.3; data not shown). However, we indeed identified a number of up regulated genes that are know to be involved in neuronal development by manual screening, eg: Hes6 (Bae et al 2010), Zfp238 (Batbet et al, 2012 and Okado et al, 2009) and Atoh8 (Inoue et al, 2001). These suggest ectopic expression of Ascl1 is able to induce the activation of neuronal specific genes at early stage of MEFs reprogramming.

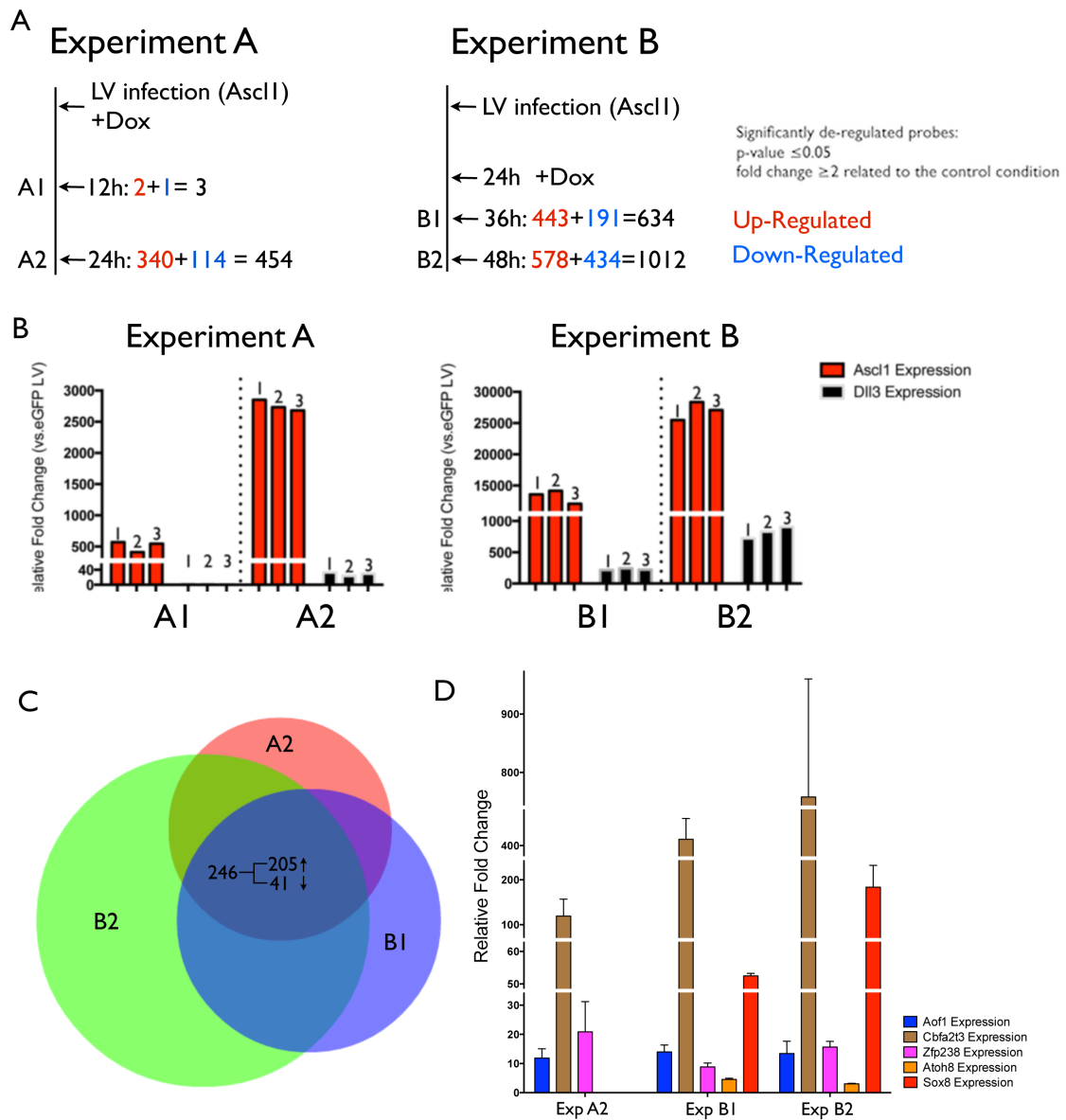


Figure 3.2: Identification of Ascl1 downstream genes using transcriptome-wide expression microarrays

Experimental setups for transcriptome-wide expression microarrays to identify Ascl1 downstream genes at the early stage of MEFs reprogramming (A). The expression of Ascl1 and Dll3 is validated using single gene qPCR as quality control. RNA samples used are collected from three biological replicates (B). Numbers of SD genes (p-value ≤ 0.05 and fold change ≥ 2) in individual experiment (A). The number of commonly deregulated genes between experiment A2, B1 and B2 (C). The expression microarray data are validated by single gene qPCR. All genes validated show significant up regulation under the Ascl1 GoF condition (D, see main text for details).

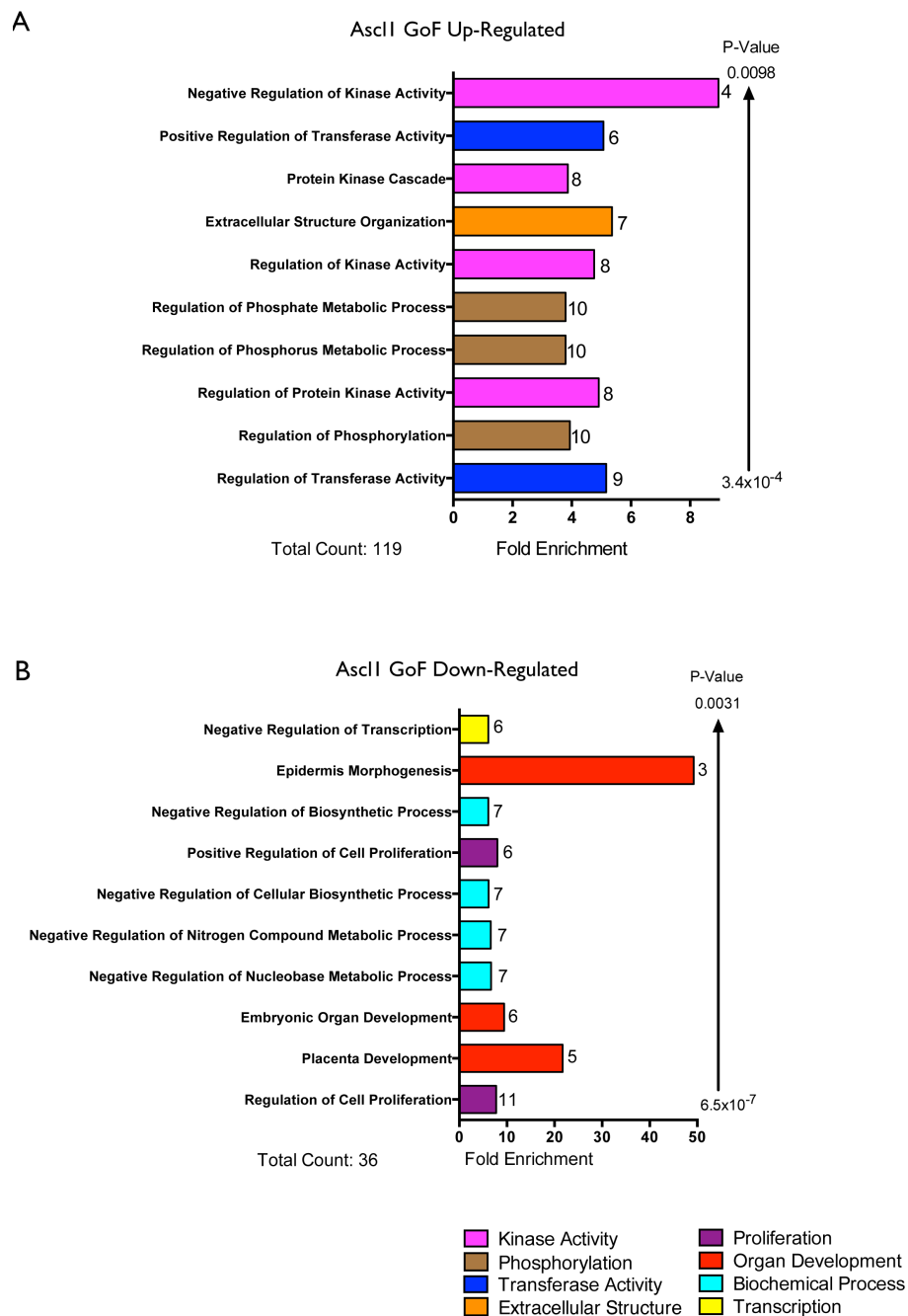


Figure 3.3: Gene ontology (GO) analysis of Ascl1 downstream genes

Enrichment of Gene Ontology biological process terms in Ascl1 downstream genes at early stage of MEFs reprogramming. Number of genes in each category is shown. Functional clusters for up-regulated (A) and down-regulated genes (B) were manually grouped based on boarder GO terms.

3.4 Selection of Ascl1 downstream gene for further study

At early stages of reprogramming, a large number of genes were deregulated upon forced expression of Ascl1. We decided to focus on 205 genes that were significantly up regulated, which were potentially involved in the activation of neuronal specific program(s) in MEFs. Known functions of these genes were manually checked (<http://www.uniprot.org/>). As mentioned earlier, we were focusing on identifying common target genes of Ascl1 involved in both MEF reprogramming and NS cell differentiation. Taking the advantage that genome wide expression profile for Ascl1-induced NS5 cell differentiation has been generated previously (Ben Martynoga, unpublished data). By cross-referencing the two data sets, we cherry picked 3 potential candidates (Zfp238, Sox8 and Atoh8) from the 70 commonly up regulated genes in both programmes (Figure 3.4). The up regulation of the candidate genes was confirmed using q-PCR (Figure 3.2 D).

Zfp238 positively regulates MyoD-mediated skeletal myogenesis (Yokoyama et al, 2009) and is essential for granule neuron precursors (GNPs) differentiation and cell cycle regulation (Tatard et al, 2010). Recent studies have also shown Zfp238 is essential for cortical neuron development (Batbet et al, 2012). Atoh8 is a bHLH TF that promotes neurogenesis while inhibit gliogenesis in developing retina (Inoue et al, 2001). Sox8 is a member of the SoxE group. Sox9 and Sox10 are the other two members of the group and have been extensively studied in the past years to show their critical functions in specification and maintenance of stem cells (Scott et al, 2010; Kim et al, 2003; reviewed by Sarkar and Hochedlinger, 2012). However, the function of Sox8 has not been well established.

Following the selection of candidates, we decided to use both GoF and LoF approaches to assess whether a candidate gene would affect Ascl1-mediated reprogramming. For the rest of the chapter, I am going to focus on Sox8.

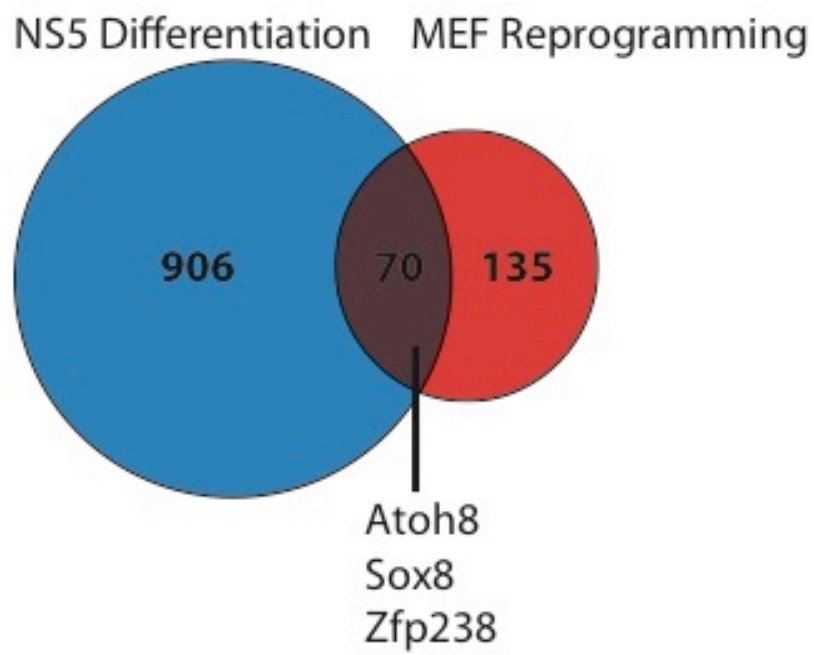


Figure 3.4: Identification of *Ascl1* downstream genes involved in both NSC differentiation and MEF reprogramming

3.5 Sox8 is involved in Ascl1-induced MEFs reprogramming

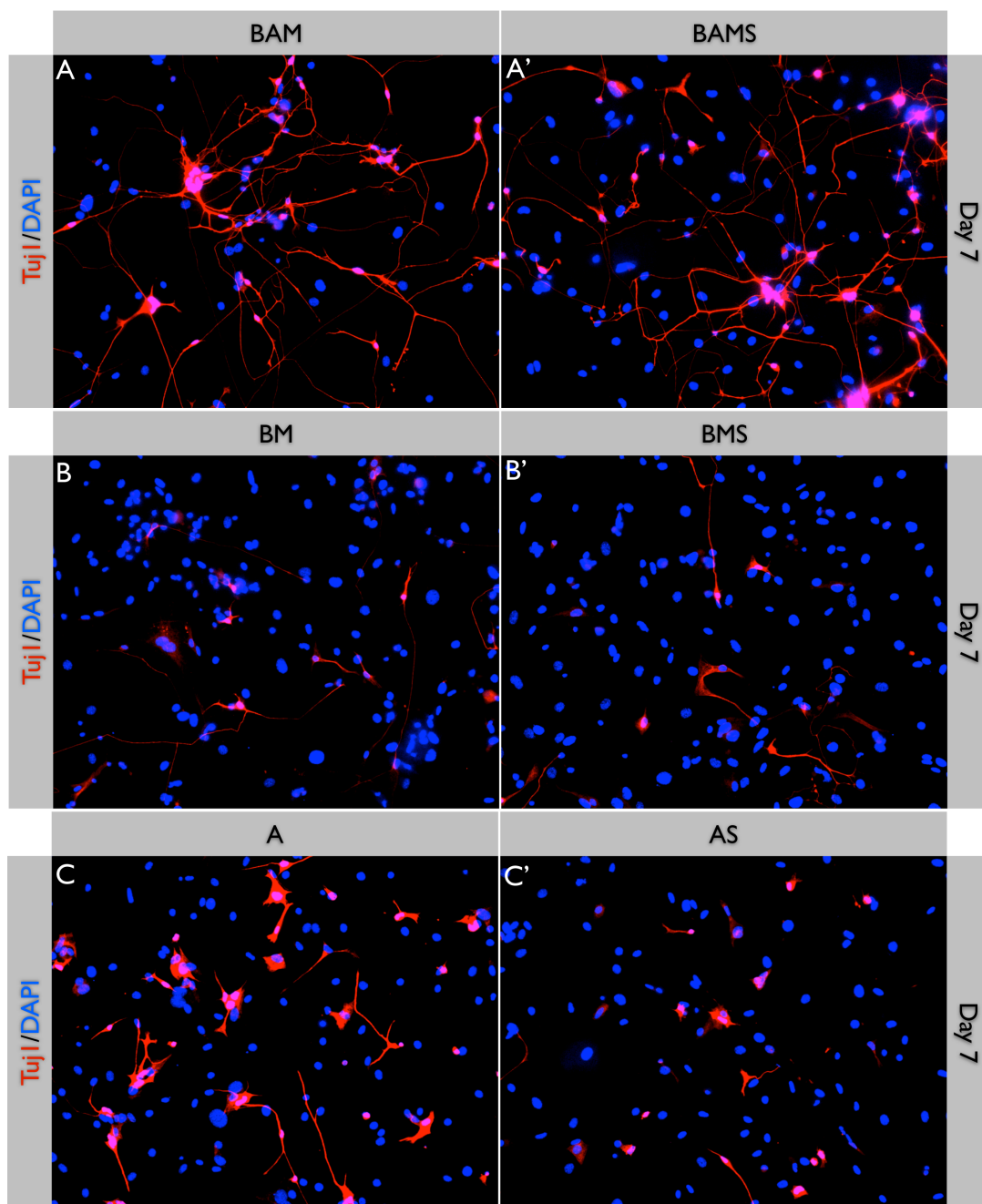
Sox8 is up regulated at the early stage of Ascl1-induced MEF reprogramming, the first questions we asked were: whether up regulation of Sox8 was essential in the reprogramming process; could Sox8 replace Ascl1 to induce efficient MEF reprogramming or enhance Ascl1-induced reprogramming efficiency. To address these questions, Sox8 GoF and LoF studies were performed.

3.5.1 Sox8 GoF reduces the generation of induced neurons (iNs)

To assess whether Sox8 could replace Ascl1 or enhance Ascl1 activity during MEF reprogramming, Sox8 was co-expressed under BAM, BM and Ascl1 only conditions. Antibody against neuronal class III β -tubulin (Tuj1) was used to detect the generation of iN at day 7 and day 13 post-infection.

At day 7, co-expression of Sox8 in the BAM cultures led to a 40% increase in number of immature neurons. However, it had no effect on the generation of iN and the reprogramming efficiency (Figure 3.5 A, A', D and D'). These suggest that ectopic expression of Sox8 cannot enhance Ascl1-mediated MEF reprogramming efficiency. Under the BM culture condition, ectopic expression of Sox8 had little effect on the generation of immature neurons and iN, which indicates that Sox8 cannot replace Ascl1 during the reprogramming process (Figure 3.5 B, B', D and D'). Under the Ascl1 only condition, ectopic expression of Sox8 had no clear effect on the generation of immature neurons but reduced the number of iNs. The reprogramming efficiency was reduced from 10% to 4% (Figure 3.5 C, C', D and D'). These suggest the possibility that Sox8 is involved in generating mature neurons from MEFs. Furthermore, no Tuj1-positive cells were observed under the Sox8 only condition (Figure 3.5 D, data not shown). This suggests Sox8 is unable to activate neuronal specific programs in MEFs. Similar results were observed at day 13 cultures (Figure 3.5 E and E', data not shown).

In conclusion, our GoF studies have shown that Sox8 cannot enhance reprogramming efficiency under the 3-TF condition and is unable to replace Ascl1 during the reprogramming process. Furthermore, Sox8 is potentially involved in generation of mature neurons. These will be discussed further in a later section.



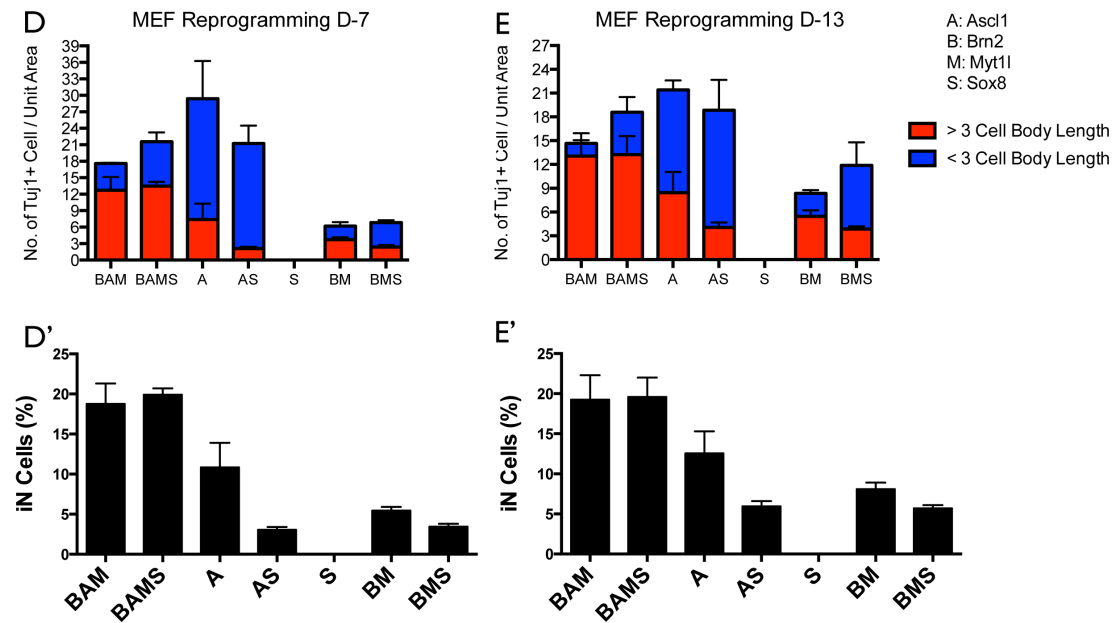


Figure 3.5: Sox8 reduces the generation of MEFs iNs

P3 MEFs were co-infected with Sox8 viruses in combination with BAM, BM and Ascl1 alone (A-C). Sox8 expression has no effect on the generation of MEFs iNs under BAM (A) and BM (B) conditions 7 days after infection. Co-expression of Sox8 with Ascl1 alone reduces the generation of iNs 7 days after infection (C). The quantification graphs show the number of Tuj1-positive cells/unit area under 20x objective and the reprogramming efficiency (calculated using only iNs) when Sox8 is co-expressed under BAM, BM and Ascl1 alone conditions 7 days (D and D') and 13 days (E and E') after infection. Sox8 alone is unable to generate Tuj1-positive cells 7 days (D and D') and 13 days (E and E') after infection.

3.5.2 Sox8 LoF reduces Ascl1-mediated generation of neuronal cells from MEFs

shRNA-mediated knock down approach was used to study whether Sox8 activity was essential for Ascl1-mediated MEFs reprogramming. shRNAs against Sox8 were designed using Whitehead siRNA (<http://sirna.wi.mit.edu/>) and cloned into viral vector contains a GFP reporter. MEFs were co-infected with lentiviruses carrying Ascl1 transgene (Sox8 is not endogenously expressed in MEFs and Ascl1 was used to induce Sox8 expression) and Sox8 shRNA. Dox was added 24h post-infection to induce Ascl1 expression and the knock down efficiency was assessed using q-PCR 72h post-infection (Figure 3.6 A). The most efficient construct was used for Sox8 LoF studies, which showed a 60% reduction of Sox8 expression at the RNA level (Figure 3.6 B).

To assess whether Sox8 activity was essential for Ascl1-induced MEF reprogramming. MEFs were co-infected with lentiviruses carrying Sox8 shRNA and BAM TFs or Ascl1 alone and analysed at day 7 and day 13 post-infection. At day 7, no clear changes in cell morphology and reprogramming efficiency were observed under the BAM condition (Figure 3.7 A, A', E and E'). Under the Ascl1 only condition, co-infection with Sox8 shRNA reduced the number of immature neurons by 50%, but had no effects on reprogramming efficiency (Figure 3.7 B, B', E and E'). At day 13, Sox8 LoF reduced reprogramming efficiency under both BAM and Ascl1 only conditions, but had no clear effect on the generation of immature neurons in both conditions (Figure 3.7 C, C', D, D', F and F').

In conclusion, LoF studies have shown that Sox8 activity is required for Ascl1-induced MEFs reprogramming. These will be discussed further in a later section.

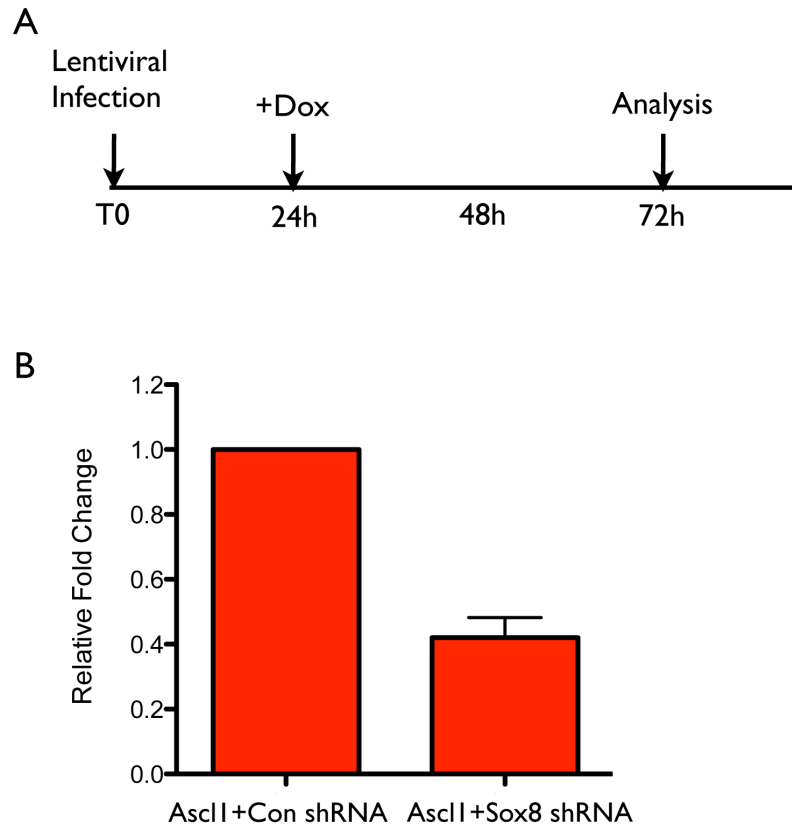
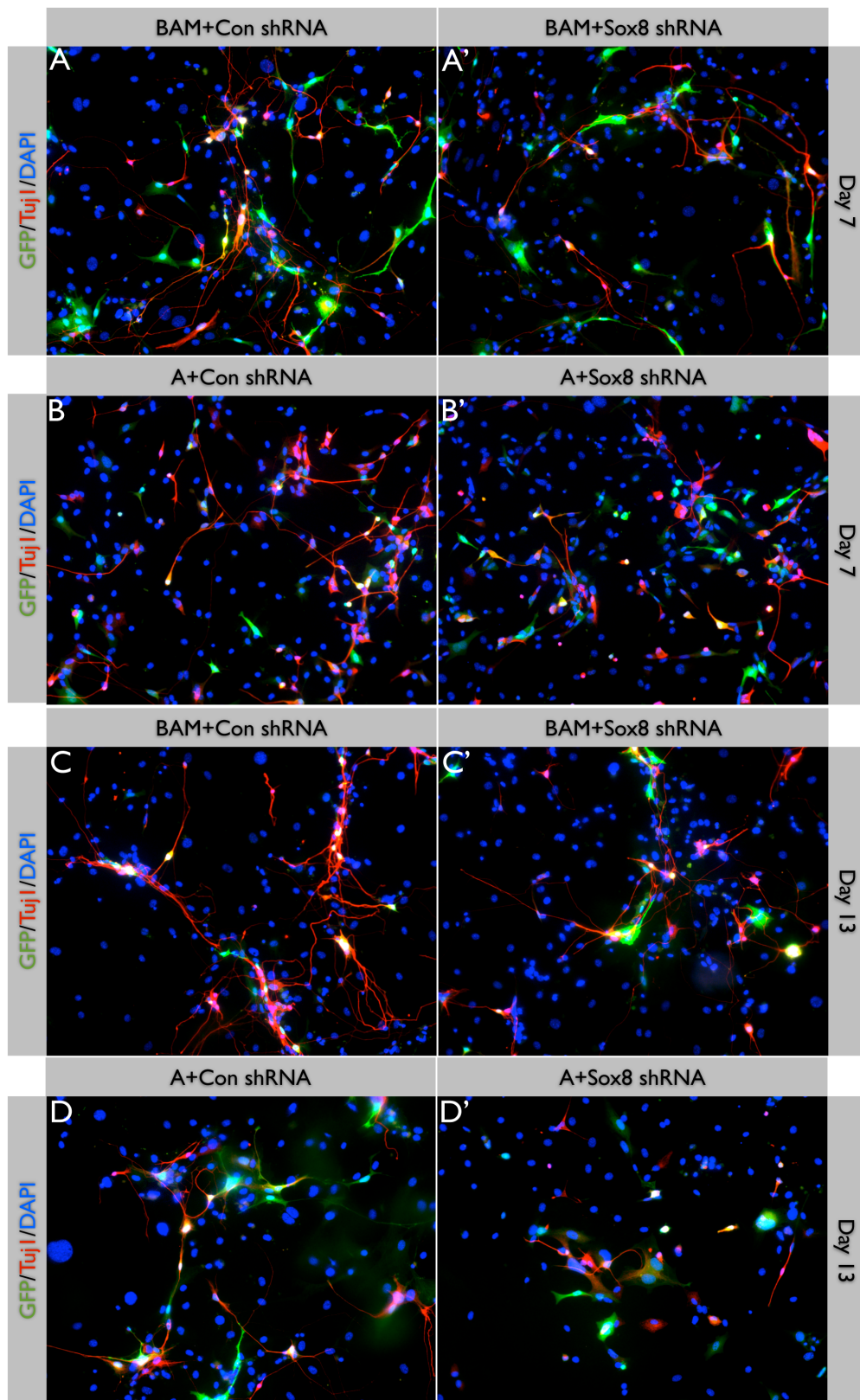


Figure 3.6: shRNA-mediated Sox8 knock down

Experimental design to test Sox8 shRNA knock down efficiency (A). P3 MEFs are co-infected with lentivirus carrying Sox8 shRNA and Ascl1. shRNA against Sox8 reduces Ascl1-induced Sox8 up regulation by 60% three days post-infection (B).



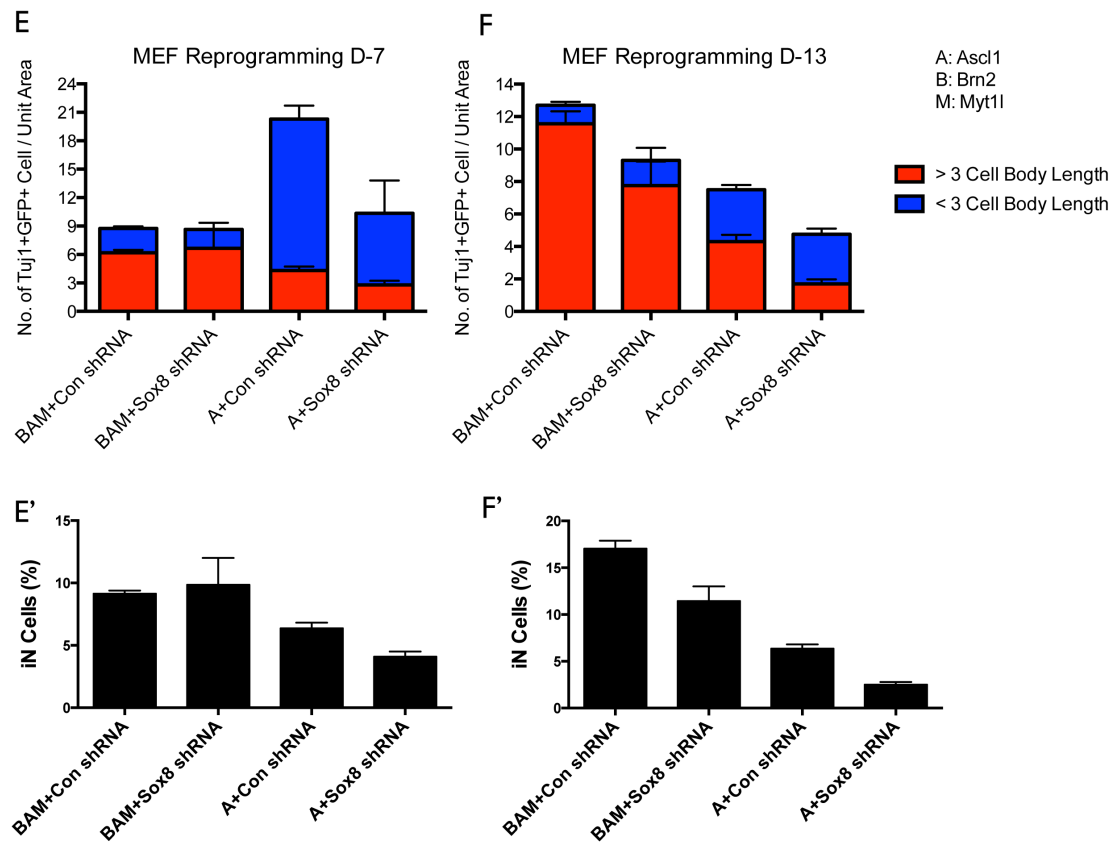


Figure 3.7: Sox8 is required for the generation of neuronal cell from MEFs

Sox8 shRNA is cloned into viral vector contains a GFP reporter. P3 MEFs were co-infected with Sox8 shRNA viruses in combination with BAM and Ascl1 alone (A-D). Sox8 LoF has no effect on the generation of MEFs iNs under BAM (A and A') and reduced the generation of immature neurons under Ascl1 alone condition (B and B') 7 days after infection. Sox8 LoF reduces the generation of iNs under both BAM and Ascl1 alone conditions 13 days after infection (C-D). The quantification graphs show the number of GFP-, Tuj1-double positive cells/unit area under 20x objective and the reprogramming efficiency (calculated using only iNs) under BAM and Ascl1 alone conditions 7 days (E and E') and 13 days (F and F') after infection.

3.6 Discussion

In this chapter, I have shown that we successfully re-established the *in vitro* reprogramming protocol using MEFs as starting cells. The MEFs iNs have classical neuronal morphology (cells have bi-polar or multi-polar structure with long extended processes) and show strong immunoreactivity for early neuronal marker Tuj1. The combination of the three neurogenic TFs (Ascl1, Brn2 and Myt1l) has the reprogramming efficiency between 12%-19%, which is comparable with previous experiments (Vierbuchen et al, 2010). Here, I will discuss our major finding, Sox8, a member of the SoxE family, is involved in Ascl1-induced MEFs reprogramming.

3.6.1 Sox8 GoF reduces the generation of MEFs iNs.

Sox8 is up regulated at the early stage of Ascl1-induced MEFs reprogramming. However, ectopic expression of Sox8 reduces the generation of MEFs iN under Ascl1 alone condition 7 and 13 days after infection (Figure 3.5 C-E', data not shown), which suggest the correct level or/and timing of Sox8 expression is required to generate iNs. Furthermore, mass spectrometry analysis has shown that Sox8 directly interact with Ascl1 at the protein level (Debbie van den Berg, unpublished data). Luciferase reporter assays indicate that Sox8 can act either cooperatively with or inhibitory to Ascl1 activity in enhancer dependent manner (Cristina Minieri, unpublished data). These suggest the possibility that Sox8 regulates Ascl1-induced activation of neuronal specific genes.

On the other hand, we are using neuronal morphologies to define iNs (cells have processes longer than three cell body length). The generation of immature neurons (cells have processes shorter than three cell body length) is not effected by Sox8 overexpression (Figure 3.5 D and E). These suggest the possibility Sox8 is involved in cytoskeletal arrangement during the generation of iNs. At the early stage of MEFs reprogramming, a number of Ascl1 downstream genes that are up regulated have cytoskeletal remodelling activities. For example, Dpysl4 is up regulated upon forced Ascl1 expression in MEFs (up regulation is confirmed by q-PCR, data not shown here) and it is involved in Sema3 signalling and subsequently cytoskeletal remodelling (Goshima et al

1995). These suggest the possibility Sox8 regulates Ascl1-induced activation of cytoskeletal remodelling genes.

Furthermore, Sox8 GoF is unable to generate Tuj1-positive cells nor enhance Ascl1-mediated reprogramming efficiency (Figure 3.5 D-E'), which suggests, unlike Ascl1, Sox8 or/and its downstream genes do not have chromatin remodelling activities to open chromatin structures during direct reprogramming of MEFs (also see 3.3.2 for further discussion).

3.6.2 Sox8 LoF reduces the generation of neuronal cells from MEFs

Sox8 LoF reduces the generation of immature neurons at day 7 post-infection under the Ascl1 alone condition, but has no noticeable effect on the mature neuron population in both efficiency and cell morphologies (Figure 3.7 E and E'). At day 13, the reduction in iNs can be partly explained by the reduced immature neuronal population at day 7 (Figure 3.7 F and F'). However, the overall reduction in the generation of Tuj1-positive cells at both time points suggests that Sox8 is required for Ascl1-mediated activation of the neuronal program in MEFs. Recent studies have shown that Ascl1 functions as a "pioneer factor" in the direct reprogramming process. ChIP-Seq experiment has shown that Ascl1 binding potentially opens the chromatin structure and allows other TFs (such as Brn2) to bind at the correct positions at the early stage of reprogramming (Wapkinski et al, 2013). These suggest that Sox8 is required, alone or together with Ascl1, to regulate a set of genes involved in reprogramming mechanisms. Whether Sox8 is involved in the Ascl1-mediated genome-wide chromatin structural change is unknown. However, Sox8 does not seem to have chromatin remodelling activities, as ectopic expression of Sox8 alone is unable to generate Tuj1-positive cells. (also see 3.3.1).

Furthermore, the additional expression of Brn2 and Myt1l seems to compensate Sox8 LoF at the early stage (Figure 3.7 E and E'). This may due to that Brn2 or/and Myt1l share common target genes with Sox8, which compensates Sox8 activity at the early stage. However, the mechanism remains unclear.

3.7 Conclusion and proposed model

Our GoF and LoF studies have shown that Sox8 is involved in Ascl1-induced neuronal program activation during MEFs reprogramming. Preliminary studies suggest the possibility that this is due to Sox8 affects Ascl1 activity in regulating genes involved in neuronal development. Furthermore, as Sox8 is indeed up regulated at early stage of the reprogramming process (Figure 3.2 D) and both GoF and LoF of Sox8 reduce the generation of neuronal cells from MEFs, it is possible that the expression of Sox8 is transient and it needs to be dissolved to allow the reprogramming process to complete (Figure 3.8). This model may partly explain that Sox8 GoF and LoF have similar phenotypes. However, at which stage of the reprogramming process Sox8 needs to be dissolved is unclear.

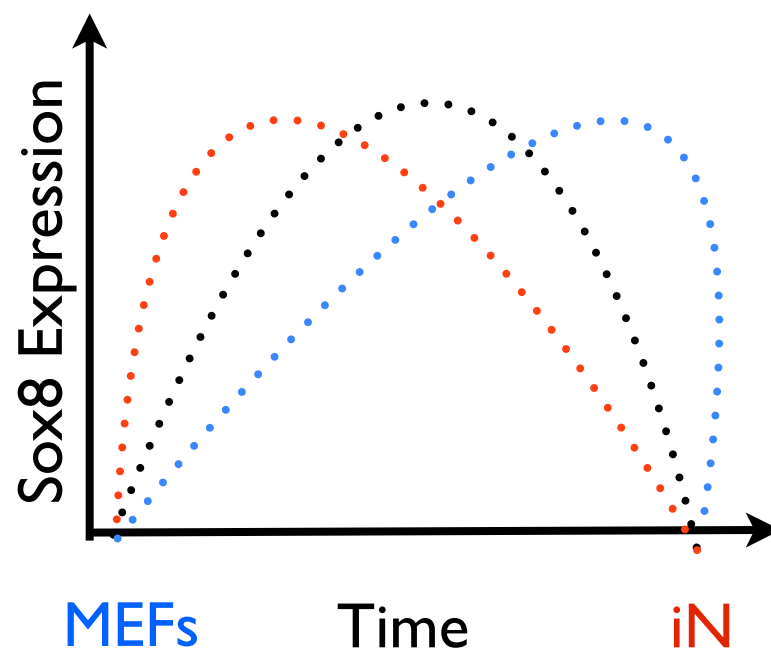


Figure 3.8: Proposed model

Sox8 is transiently expressed during the generation MEFs iN. Sox8 needs to be dissolved at either early (red dotted line), middle (black dotted line) or late (blue dotted line) stage of the reprogramming process.

Chapter 4: Functions of SoxE in neural stem cell maintenance *in vitro*

4.1 Rational, hypothesis and aim

Pervious studies have shown that *Ascl1* regulates genes involved in all steps of neurogenesis in a direct manner (Castro et al, 2011). *In situ* hybridization data have shown that Sox8 is highly expression in ventricular zone (VZ) and subventricular zone (SVZ) in the developing mouse telencephalon, which contain the proliferating progenitor population (Figure 4.1 A). The expression pattern is partly overlapped with *Ascl1* (Figure 4.1 C). I have shown that Sox8 is involved in the activation of *Ascl1*-induced neuronal specific program during MEFs reprogramming. Genome wide expression profile indicates that Sox8 is up regulated during *Ascl1*-induced NS5 cell differentiation (Ben Martynoga, unpublished data). Furthermore, preliminary data have shown that Sox8 directly interact with *Ascl1* at the protein level (Debbie van den Berg, unpublished data) and acts either cooperatively with or inhibitory to *Ascl1* in an enhancer dependent manner (Cristina Minieri, unpublished data). These suggest Sox8 is involved in *Ascl1*-induced neuronal differentiation.

Furthermore, *In situ* hybridization data have also shown that Sox9, another member of the SoxE group, has the similar expression pattern as Sox8 (Figure 4.1 B). In addition, GoF and LoF studies using transgenic mouse lines have show that Sox9 is involved in the generation and maintenance of neural stem cell in developing CNS (Scott et al, 2010). Also, the DNA binding domain, high mobility group (HMG), is highly conserved between Sox8 and Sox9 (Figure 4.1 D), which suggests the function redundancy between the two members of the SoxE group. These data raise the possibility that Sox8 is involved in NSC maintenance.

Based on the data mentioned above, we hypothesised that Sox8 is involved in NSC maintenance and differentiation.

Indeed, Sox9 has been shown to be involved in NSC maintenance (Scott et al, 2010). However, the molecular mechanism remains unclear. Also, due to

possible functional redundancy between the two members, I decided to include Sox9 in my studies.

To assess whether SoxE (specifically Sox8 and Sox9) was involved in NSC maintenance and differentiation, I decided to use both GoF and LoF approaches and NS5 as the *in vitro* system. As shown in earlier experiments, Sox8 GoF is involved in Ascl1-induced generation of iNs from MEFs. Using similar approach, Ascl1 and SoxE were co-expressed to assess SoxE functions in regulating Ascl1-mediated NSC differentiation. shRNA-mediated KD approach was used to assess whether SoxE was involved in maintaining NSCs in a undifferentiated proliferating state.

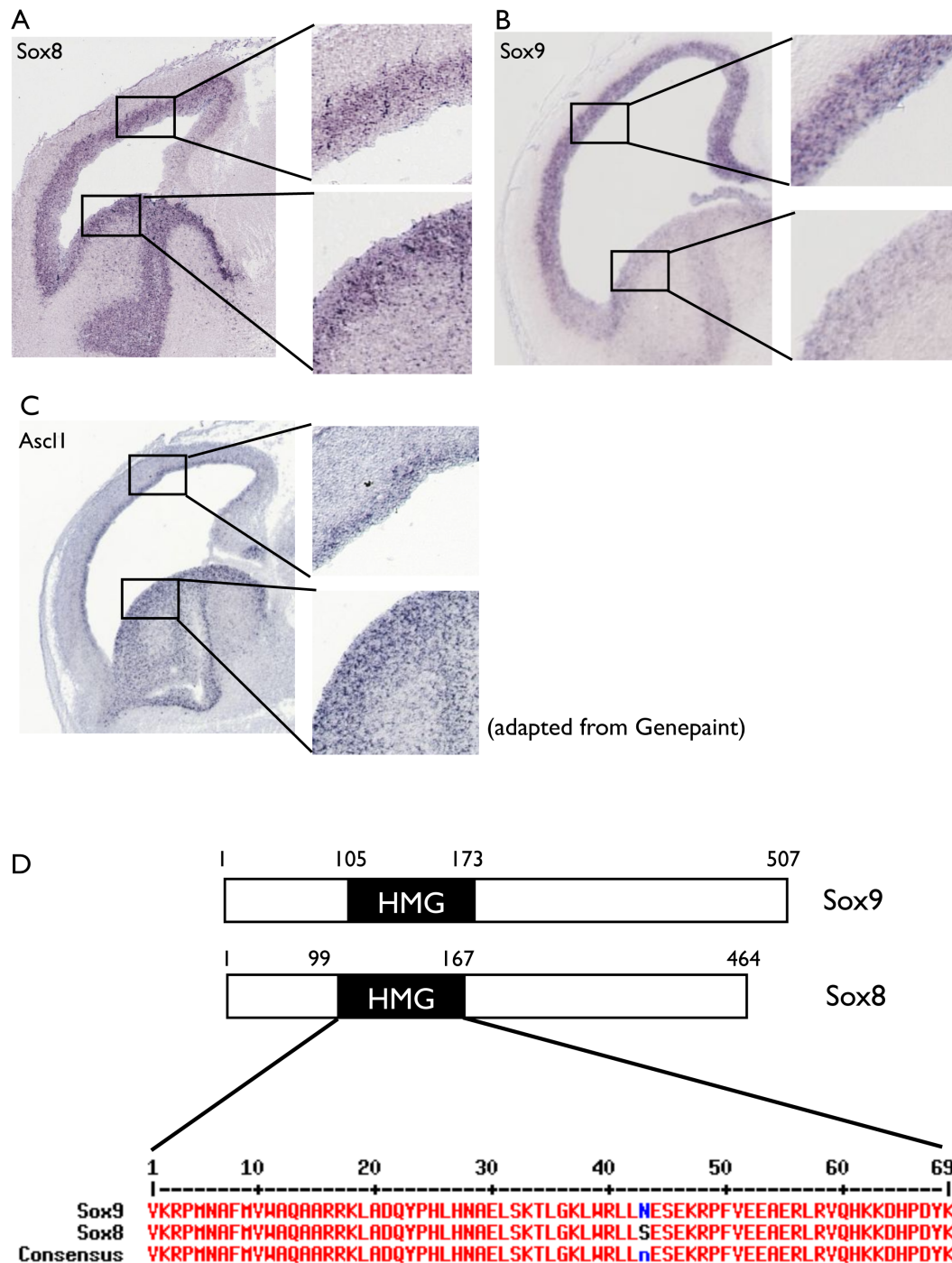


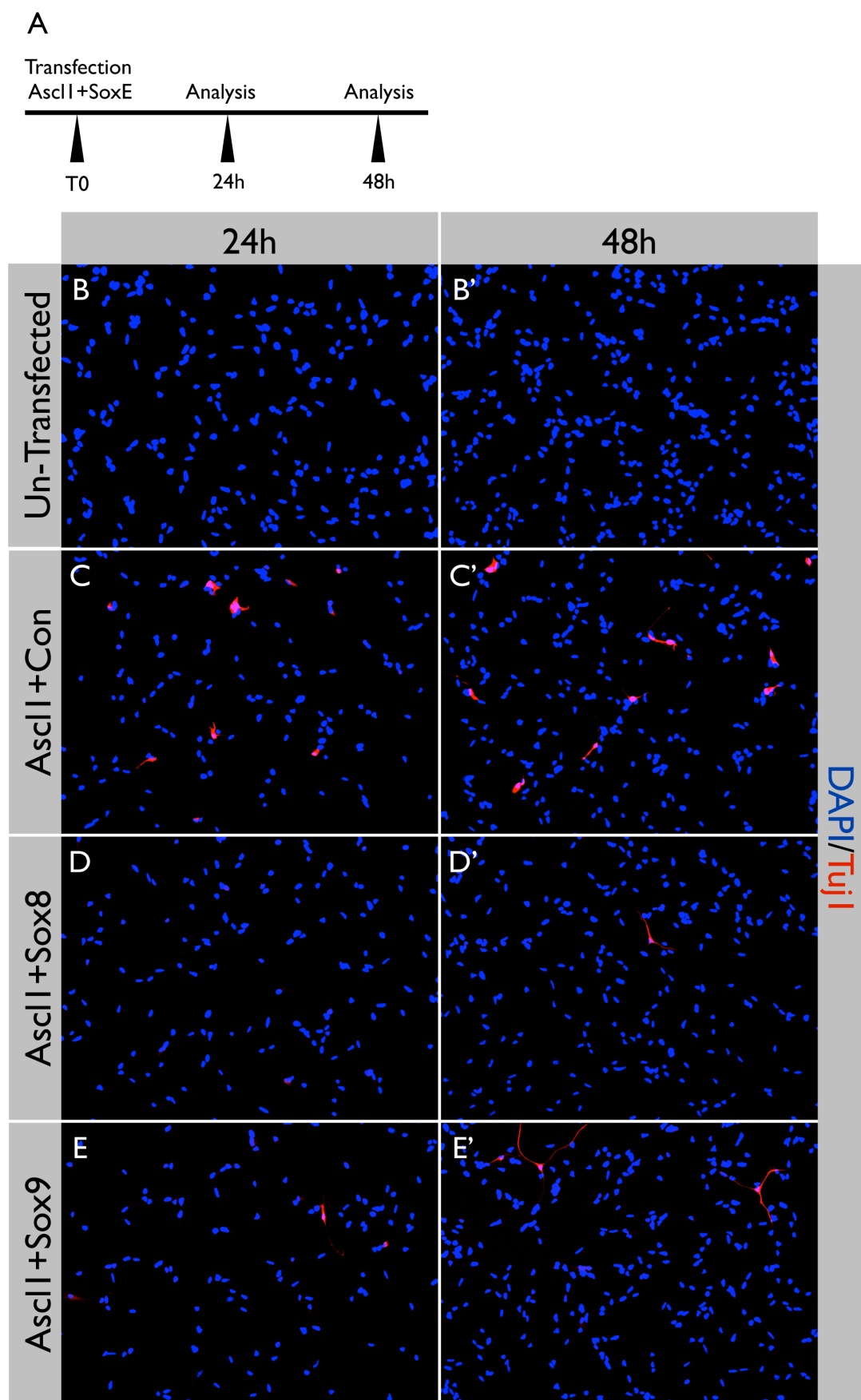
Figure 4.1: Expression of Sox8, Sox9 and Ascl1 in E14.5 mouse telencephalon and graphical view of Sox8 and Sox9 protein

In situ hybridization data to show the expression pattern of Sox8 (A), Sox9 (B) and Ascl1 (C) in E14.5 mouse telencephalon. A graphical view of Sox8 and Sox9 proteins (D). The DNA binding domain (High Mobility Group: HMG) is highly conserved and the position is highlighted.

4.2 SoxE inhibits Ascl1-induced NSC differentiation *in vitro*

Sox8 or Sox9 was co-transfected with Ascl1 to assess its function(s) in Ascl1-induced NSC differentiation. To ensure all cells transfected with Ascl1 also received Sox8 or Sox9, the ratio of the amount of DNA between Ascl1 and Sox8 or Sox9 was 1:3. Antibody against neuronal marker Tuj1 was used to identify differentiated neuronal cells 24h and 48h post-transfection (Figure 4.2 A). No Tuj1-positive cells were observed when un-transfected at both time points (Figure 4.2 B and B'). Ectopic expression of Ascl1 triggered neuronal differentiation of NSCs, which was indicated by the Tuj1 staining (Figure 4.2 C and C'). The number of Tuj1-positive cells was clearly reduced when cells were co-transfected with Sox8 (Figure 4.2 D and D') or Sox9 (Figure 4.2 E and E') at both time points. The observation was confirmed by quantification. Overexpression of Sox8 or Sox9 significantly reduced the generation of Tuj1-positive cells, which was induced by ectopic expression of Ascl1. The percentage of Tuj1-positive cells was reduced from 5% and 3% to less than 1% at 24h and 48h time points (Figure 4.2 F).

These data suggest that SoxE inhibits Ascl1-induced neuronal differentiation of NSCs *in vitro*. This will be discussed further in later section.



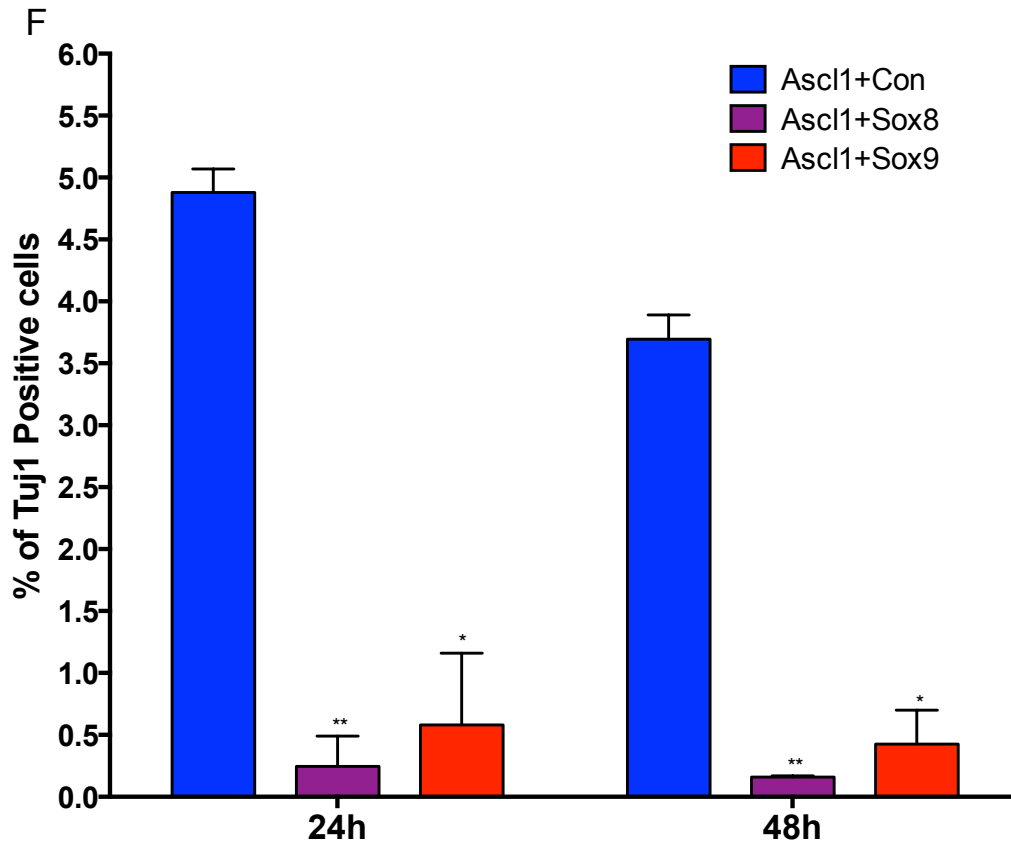


Figure 4.2: SoxE inhibits Ascl1-induced neuronal differentiation *in vitro*

Experimental rational (A). Un-transfected NS5 cells (B and B'). Ascl1 was co-transfected with control plasmid (C and C'), Sox8 (D and D') and Sox9 (E and E'). Neuronal marker Tuj1 (red) was used to identify differentiated neurons. The quantification graph shows percentage of Tuj1-positive cells 24h and 48h post-transfection (F). Data shown are mean of two independent biological replicates. Error bar: mean \pm SEM. t test: * $p < 0.05$, ** $p < 0.01$.

4.3 SoxE activity is involved in NSC proliferation *in vitro*

To assess whether SoxE is involved in NSC proliferation, I decided to perform SoxE LoF studies using shRNA-mediated KD approach. shRNAs against Sox8 or Sox9 were designed using Whitehead siRNA (<http://sirna.wi.mit.edu/>) and cloned into a plasmid contains puromycin resistance cassette for selection. q-PCR and western blot were used to examine the KD efficiency at both mRNA and protein level. Due to possible functional redundancy between Sox8 and Sox9, the Sox8-Sox9 double KD experiment was also performed to avoid functional compensation between the SoxE members.

4.3.1 shRNA-mediated KD efficiently reduces the expression of Sox8 and Sox9

Due to lack of antibody against Sox8 protein, an inducible cells line overexpressing V5-tagged Sox8 was generated from NS5 cells (iSox8-V5 NSC). Puromycine was added 16h post-transfection for selection and doxycycline was used to induce the expression of V5 tagged Sox8 protein. RNA and protein samples were collected 72h post-transfection (Figure 4.3 A). The presence of doxycycline induced the expression of V5-tagged Sox8 (Figure 4.3 A' and A''). The q-PCR analysis showed a 60% and 70% reduction of Sox8 expression at the RNA level in the single and double KD experiments (Figure 4.3 A'). The western blot experiment indicated a clear reduction of V5-tagged Sox8 protein, which was absent without doxycycline induction (Figure 4.3 A'').

NS5 cells were used to assess the KD efficiency of the Sox9 shRNA (Figure 4.3 B). Similarly, protein and RNA samples were collected 72h post transfection with 56h puromycine selection. q-PCR analysis showed a 50% and 40% reduction of Sox9 RNA under single and double KD conditions (Figure 4.3 B'). At the protein level, the expression of Sox9 was clearly reduced (Figure 4.3 B'').

In summary, shRNA against Sox8 or Sox9 efficiently reduces the expression of its target gene at both transcriptional and translational levels under both single and double KD conditions.

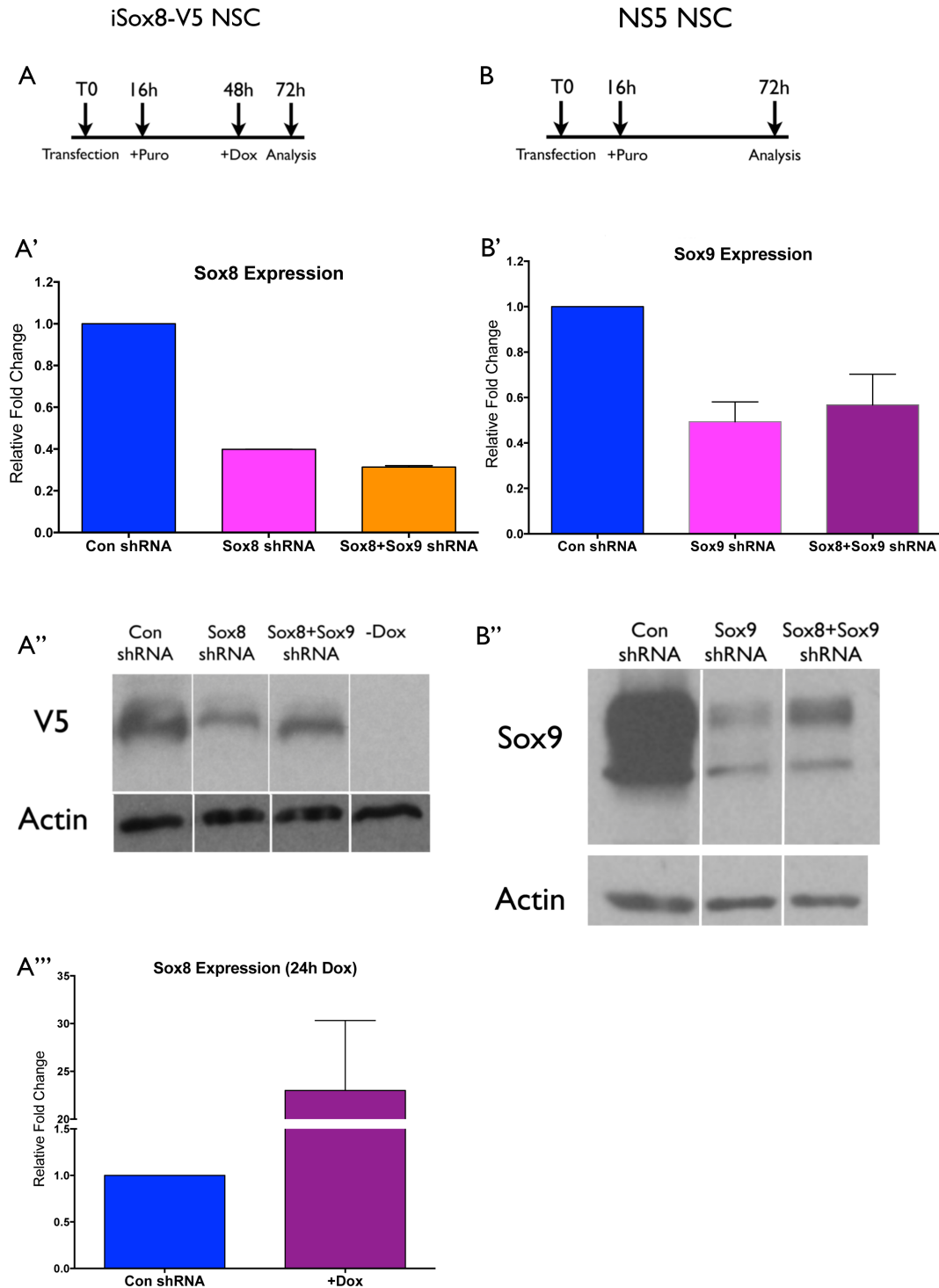


Figure 4.3: shRNA-mediated SoxE knock down

Experimental design to test shRNA knock down efficiency (A and B). q-PCR analysis to examine the reduction of Sox8 (A') and Sox9 (B') expression at the RNA level under single and double KD conditions. Western blot to assess the reduction of V5-tagged Sox8 (A'') and Sox9 (B'') protein. q-PCR analysis to indicate the up-regulated of Sox8 expression after doxycycline induction (A''').

4.3.2 SoxE is essential for NSC proliferation but not required to maintain NSC in an undifferentiated state

shRNA-mediated SoxE KD approach was used to assess whether SoxE activity was involved in NSC proliferation. shRNA against Sox8 or Sox9 was re-cloned into a vector contains a GFP reporter. EdU, which was used to measure S-phase progression, was applied 48h post-transfection for two hours (Figure 4.4 A). Under the control condition, more than 40% of the transfected cells were able to enter S-phase (Figure 4.4 B and C, indicated by EdU-GFP double positive cells). The percentage of transfect cells entered S-phase was significantly reduced to 8% and 19% under Sox8 and Sox9 LoF conditions (Figure 4.4 B', B'' and C). Under the double KD condition, it was reduced to less than 5% (Figure 4.4 B''' and C).

In addition, antibodies against Tuj1 and DCX (neuronal markers), GFAP (an astrocyte marker), O4 and NG2 (oligodendrocyte markers) were used to assess whether SoxE was involved in maintaining NSCs in an undifferentiated state. However, GFP-positive cells were not immunoactive to any of the antibodies, which was suggesting the transfected NSCs remains undifferentiated (data not shown).

The preliminary data have show that SoxE activity is essential to maintain NSC proliferation *in vitro*. Furthermore, there is no functional compensation between Sox8 and Sox9 in regulating NSC proliferation. These will be discussed in a later section.

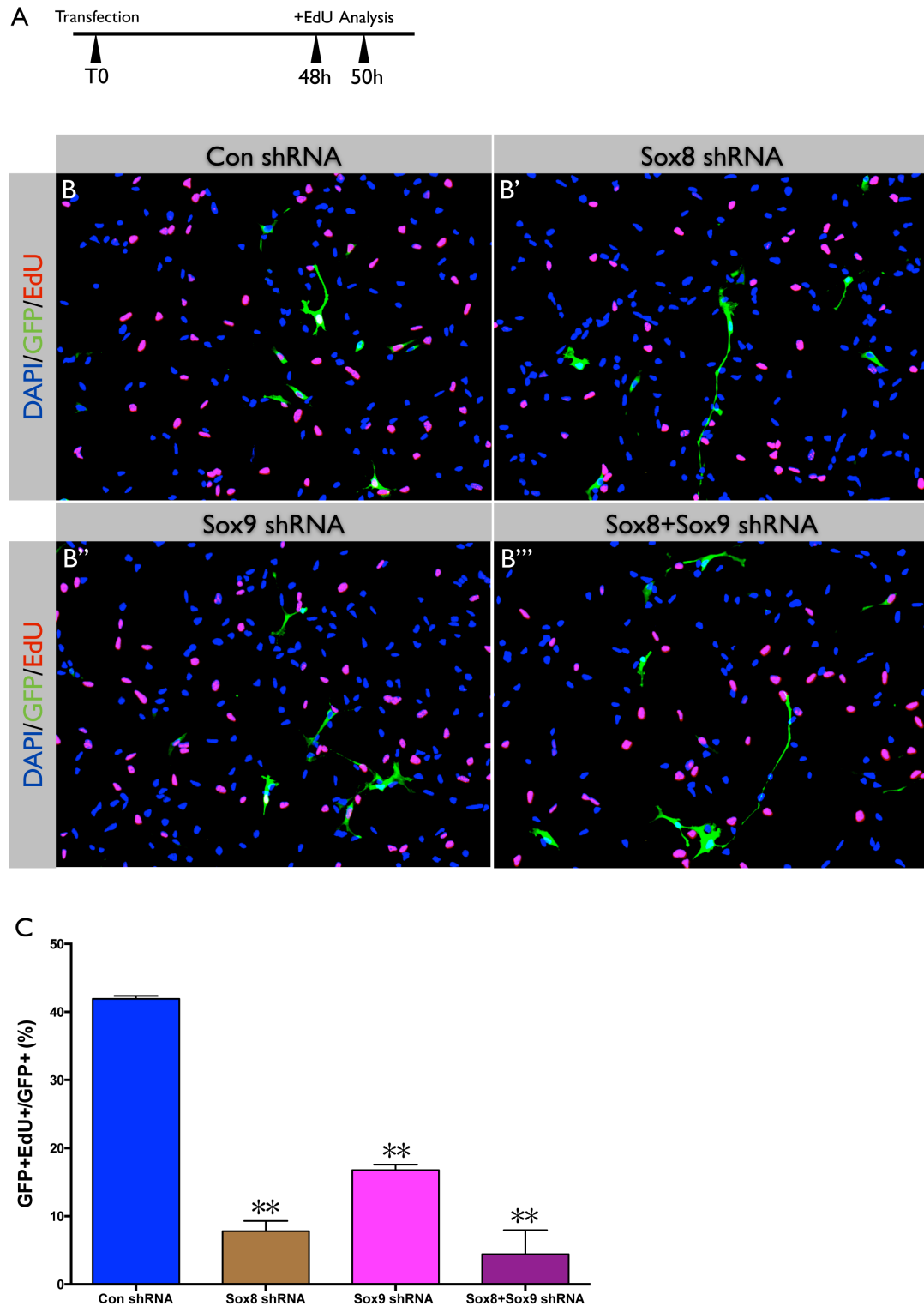


Figure 4.4: SoxE activities are essential in maintaining NSC proliferation
 Experimental design to assess SoxE functions in NSC proliferation (A). Sox8 LoF (B'), Sox9 LoF (B'') and Sox8-Sox9 double KD (B''') reduced the number of EdU positive cells (red) within the transfected population (green) comparing with control (B). The quantification graph shows percentage EdU-GFP double positive cells 50h post-transfection (C). Data shown are mean of two independent biological replicates. Error bar: mean \pm SEM. t test: **p<0.01.

4.3.3 Identification of SoxE downstream genes

Our preliminary data have shown that SoxE activity is essential to maintain NSC proliferation. In order to understand the molecular mechanism involved, I decided to use genome-wide expression microarrays to detect gene expression changes under SoxE LoF conditions. RNA samples were collected 72h post transfection with 56h puromycin selection (Figure 4.5 A) and collected from three biological replicates. As a quality control, q-PCR analysis was used to assess SoxE knock down efficiency and showed a 40%-70% reduction at the RNA level under single and double KD conditions (Figure 4.5 B). mRNA samples were hybridized to Illumina Ref8v2 BeadArrays (performed by Harsha Jani, NIMR sequencing facility) and analysed using the GeneSpring software. The minimum requirements for significantly deregulated (SD) genes were a corrected p-value lower than 0.05 and a relative fold change higher than 1.5.

In order to validate the quality of the gene expression profiling data, 12 SD genes were cherry picked (6 up-regulated and 6 down-regulated) and the expression level were validated using q-PCR. Samples used for expression microarrays and two additional biological replicates were used for q-PCR validation. The 11 out of 12 genes validated were significant deregulated under the SoxE LoF conditions, which suggested the gene expression profiles were reliable (data not shown).

The numbers of SD genes for single and double KD were shown in Figure 4.5 C. Comparing Sox8 and Sox9 single KD conditions, 520 genes were deregulated by Sox8 shRNA only (255 down regulated and 265 up regulated). 875 genes were de-regulated by Sox9 shRNA only (368 down regulated and 507 up regulated). 409 genes were commonly de-regulated by Sox8 and Sox9 shRNAs (157 down regulated and 252 up regulated). 724 genes were de-regulated under the double KD condition only (397 down regulated and 327 up regulated) (Figure 4.5 C, D and D'). This will be discussed further in a later section.

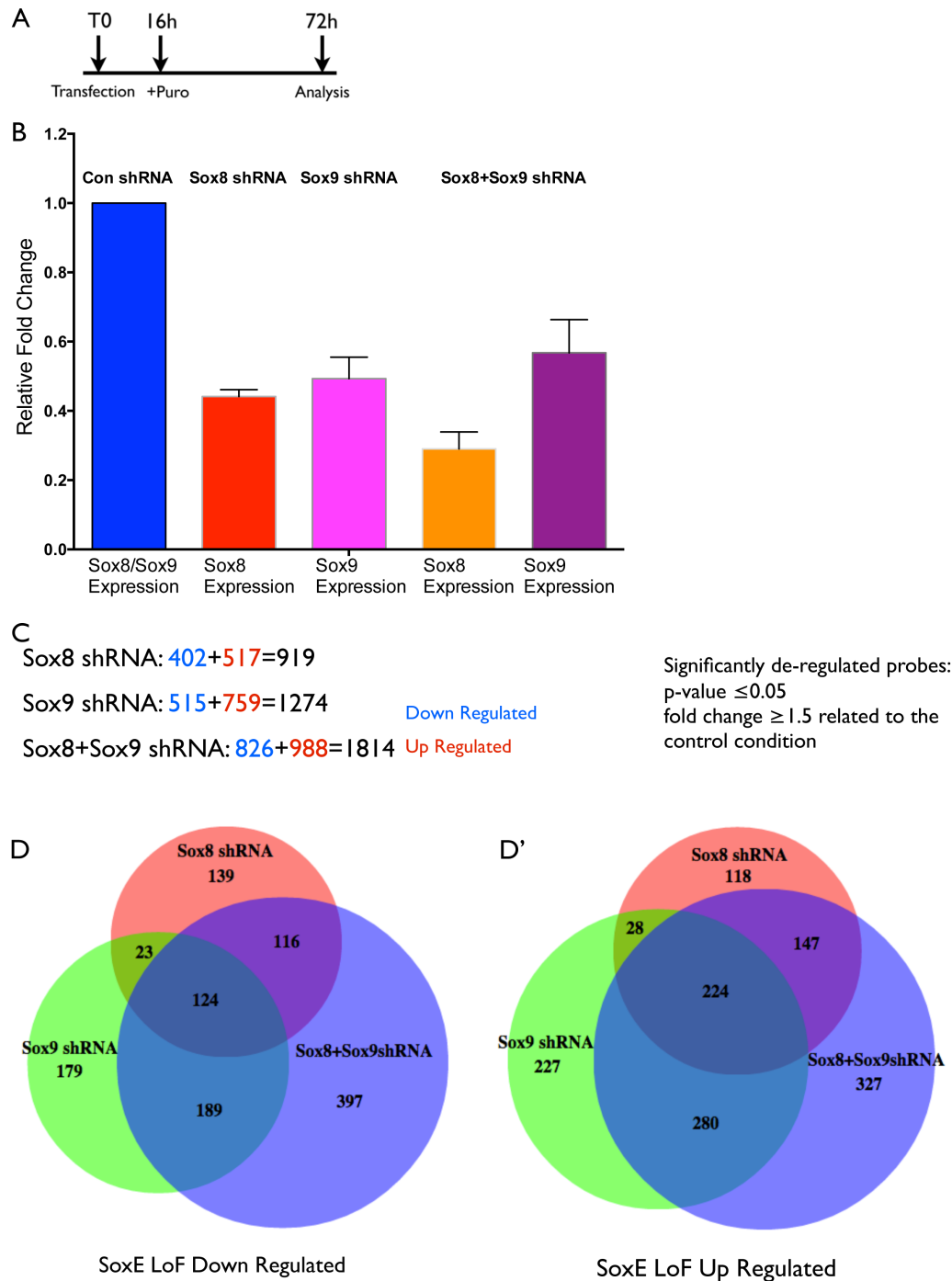


Figure 4.5: Identification of SoxE down stream genes using genome-wide expression microarrays

Experimental setups for transcriptome-wide expression microarrays (A). The expression of Sox8 or/and Sox9 under single and double knock down conditions is validated using single gene qPCR as quality control. RNA samples used are collected from three biological replicates (B). Numbers of SD genes (p-value ≤ 0.05 and fold change ≥ 1.5) under single and double KD conditions (C). Number of commonly and uniquely SD genes under single and double knock down conditions (D and D').

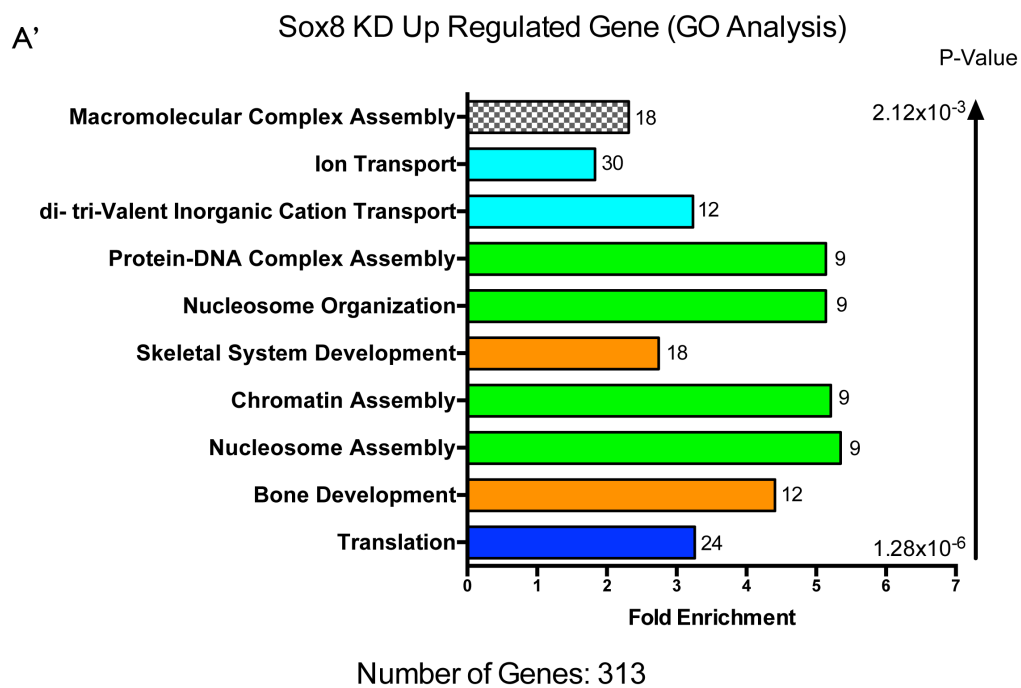
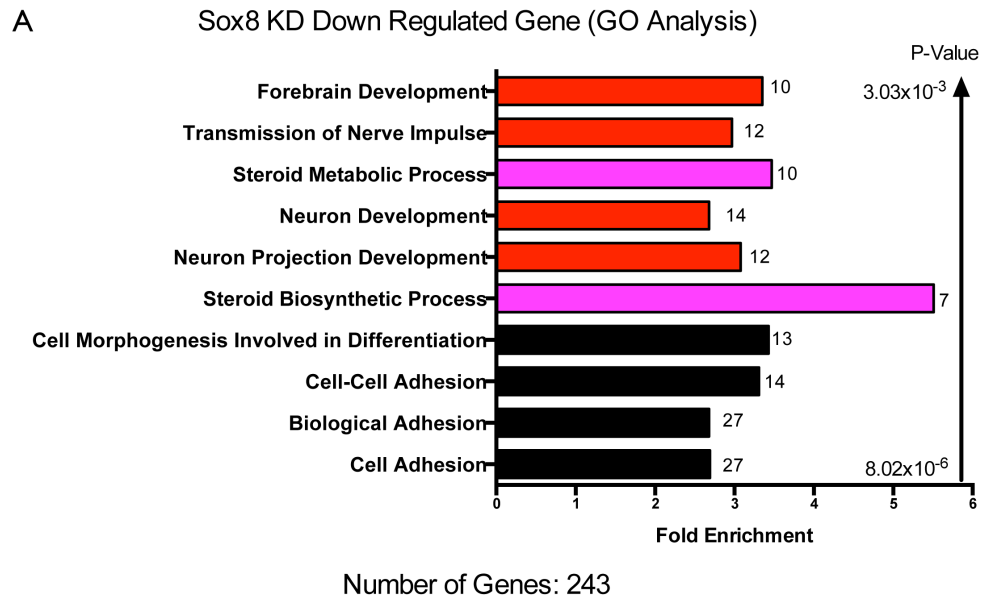
4.3.4 Functional annotation of SoxE downstream genes

To have global understanding of functions of SoxE downstream genes, SD genes were annotated to functional clusters using Gene Ontology (GO) analysis (<http://david.abcc.ncifcrf.gov/>). The top ten functional clusters for up regulated (Sox8 KD, Sox9 KD, Sox8-Sox9 double KD, Figure 4.6 A', B' and C') and down regulated genes (Figure A, B and C) were manually grouped based on boarder GO terms.

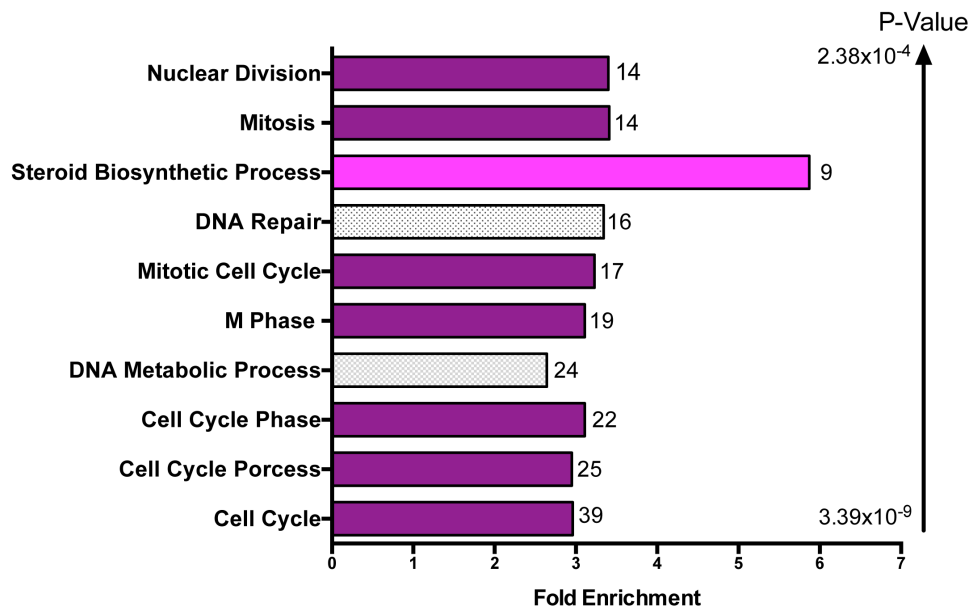
Under the Sox8 knock down condition, genes involved in cell adhesion, sterol process and neuron development were down regulated. For clusters related with neuron development, no neuronal specific genes were down regulated (eg: Ngns, SoxC and SoxB2). Genes involved in axon formation (eg: Stmn1) and neurite development (eg: Chl1 and Slit) were observed. Genes involved in bone development and chromatin organization were up regulated. Clusters related with chromatin organization were formed by histone genes such as Hist1H4M and Hist2H3C1 (Figure 4.6 A and A').

Under the Sox9 knock down condition, genes involved in cell cycle regulation were down regulated (eg: Cyclin A2, Cyclin B1 and Cdc2A). Genes related with blood vessel development, bone development and neuron differentiation were up regulated (Figure 4.6 B and B'). The cluster related with neuronal differentiation did not contain neuronal specific genes such as Ngn, SoxC or SoxB2. Ephrin receptor genes were observed in the group (EphB1 and EphA1).

Under the Sox8-Sox9 double KD condition, similar to Sox9 KD, genes related to cell cycle regulation were down regulated (cyclin family members, cell division control protein members and cell division control associated protein members). Genes that were up regulated were annotated into clusters involved in diverse functions (Figure 4.6 C and C'). The results will be discussed further later.

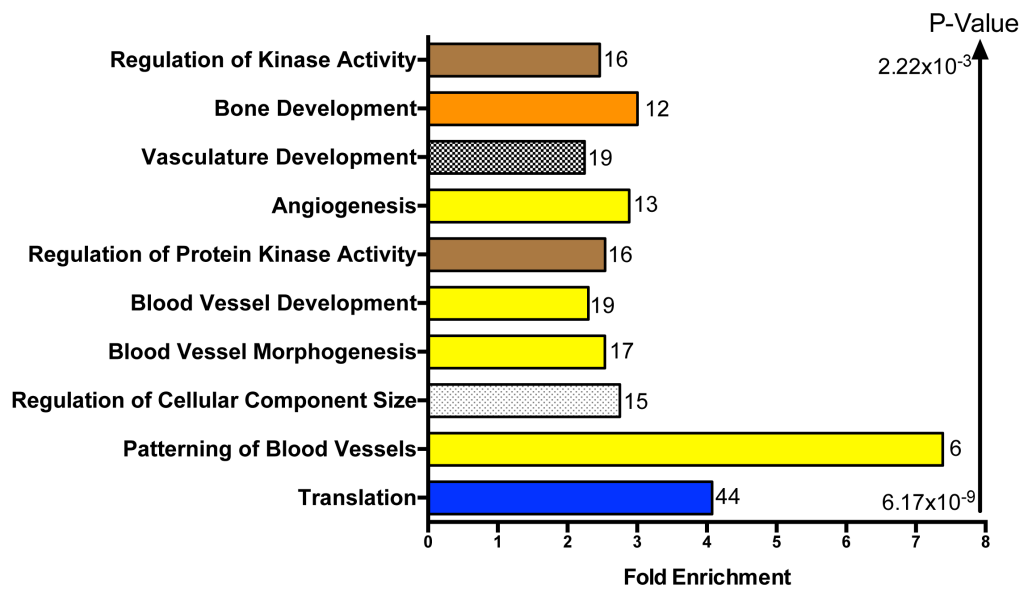


B Sox9 KD Down Regulated Gene (GO Analysis)



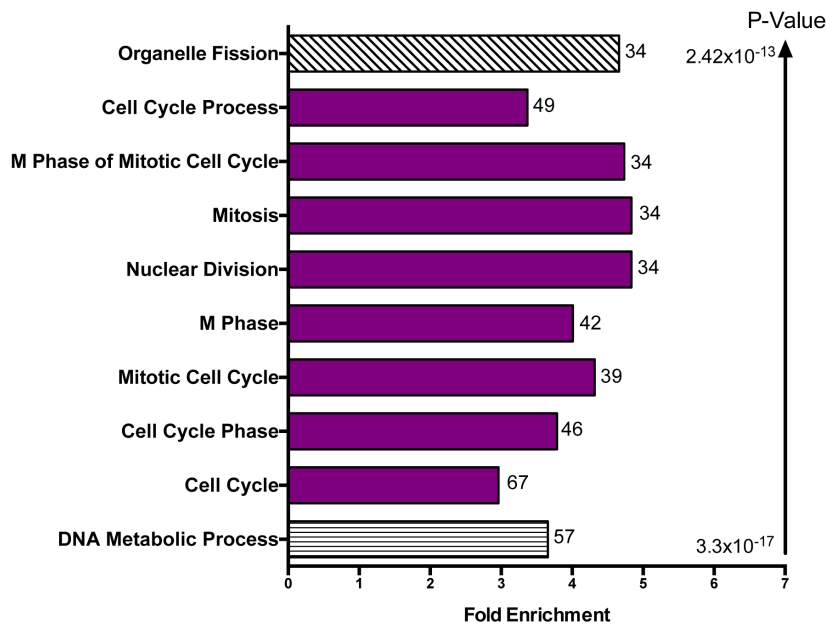
Number of Genes: 293

B' Sox9 KD Up Regulated Gene (GO Analysis)



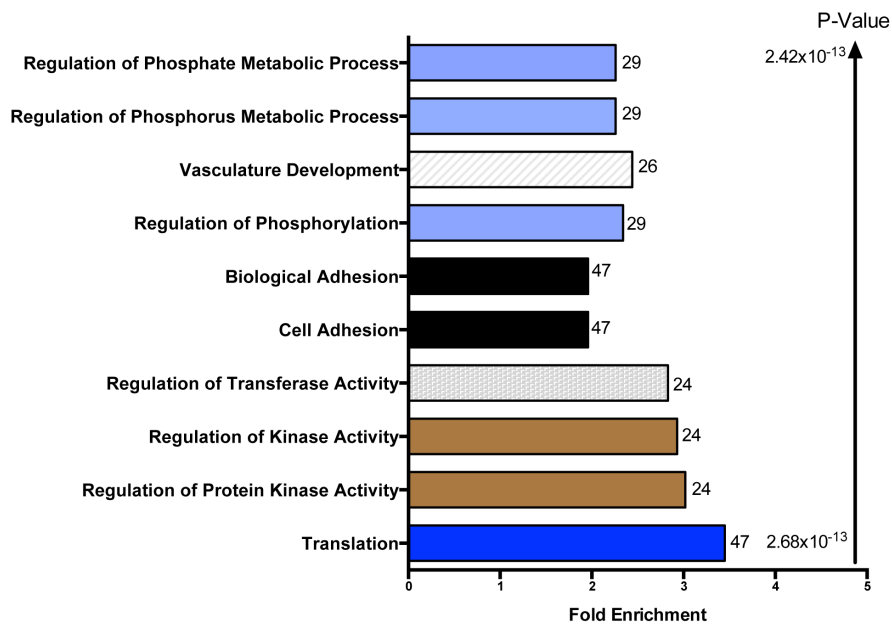
Number of Genes: 460

C Sox8+Sox9 KD Down Regulated Gene (GO Analysis)



Number of Genes: 503

C' Sox8+Sox9 KD Up Regulated Gene (GO Analysis)



Number of Genes: 579



Figure 4.6: Gene Ontology (GO) analysis of Ascl1 downstream genes

Enrichment of Gene Ontology biological process terms in SoxE downstream genes in NS5 cells. Number of genes in each category is shown. Functional clusters for down-regulated (A, B and C) and up-regulated genes (A', B' and C') were manually grouped based on boarder GO terms.

4.3.5 SoxE regulates genes involved in cell cycle regulation

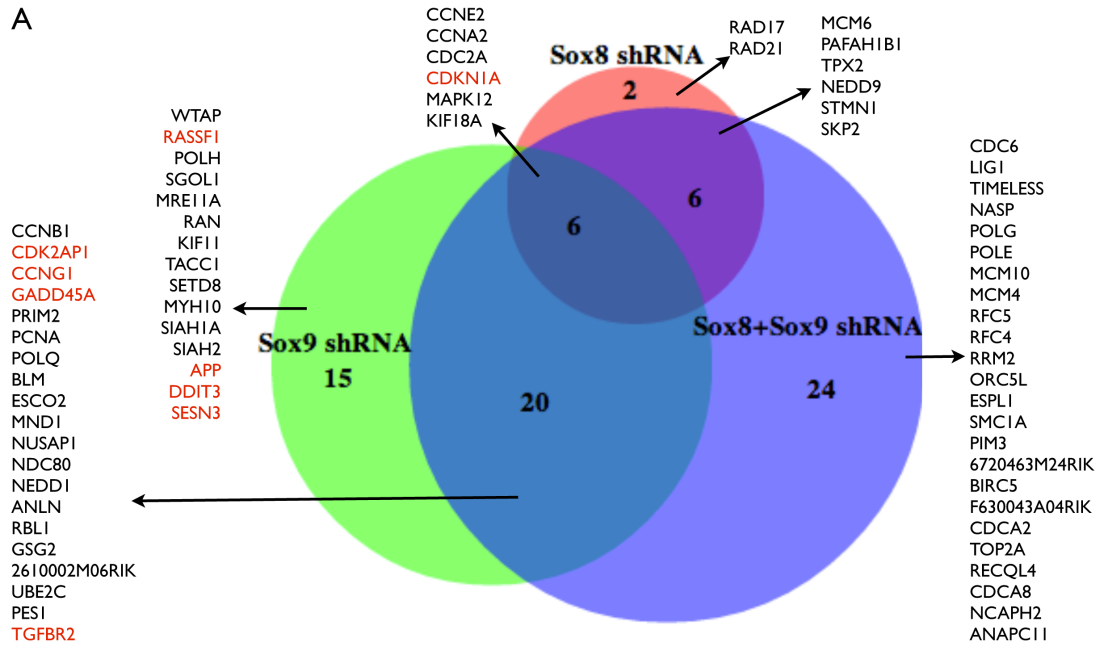
As shown in my earlier experiments, SoxE activity is potentially involved in maintaining NSC proliferation *in vitro* (Figure 4.4). Also, genome-wide expression microarray data and GO analysis have shown that SoxE downstream genes are involved in cell cycle regulation (Figure 4.6, data not shown). Here, I decided to focus on this group of genes.

Functional clusters related with cells cycle regulation were manually selected (in order to pick up all cell cycle regulation genes, the p-values of the functional clusters were not considered at this point).

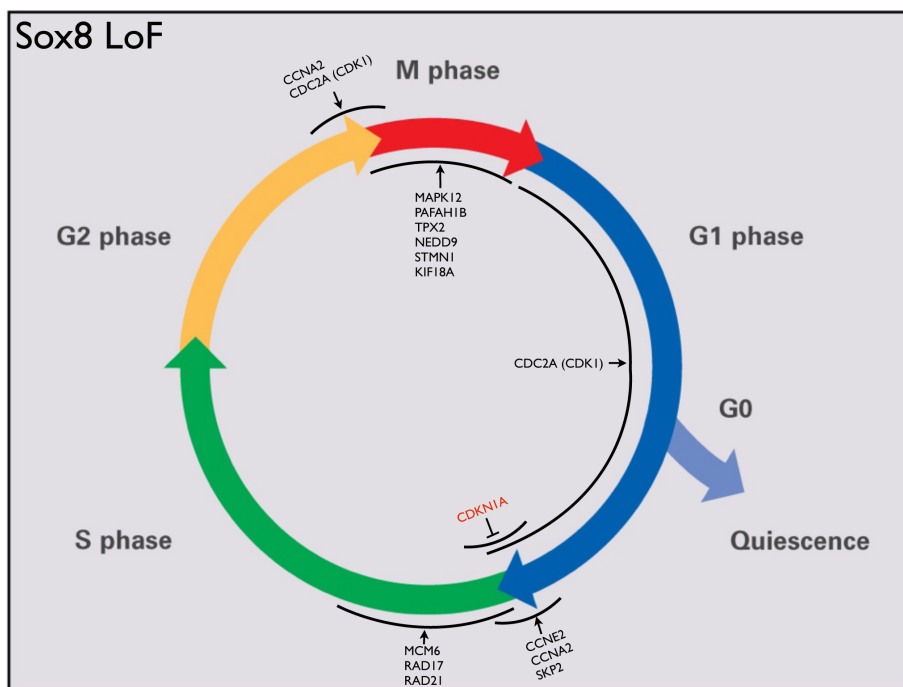
Under the Sox8 KD condition, 14 cell cycle related genes were identified (13 down-regulated and 1 up-regulated). Under the Sox9 KD condition, 41 cell cycle related genes were deregulated (32 down regulated and 9 up regulated). Under the double KD condition, 56 cell cycle related genes were regulated (51 down regulated and 5 up regulated). 6 genes were commonly deregulated by Sox8 and Sox9, but majority of the cell cycle related genes downstream of Sox8 or Sox9 were different. 24 genes were deregulated under the double KD condition, which suggested functional redundancy between the two SoxE members (Figure 4.7 A).

Individual genes were manually fitted into different phases of the cell cycle process based on the known functions (<http://www.uniprot.org/>). It was clear that SoxE downstream were related with multiple phases of the cell cycle, including phase transitions (Figure 4.7 B-D). The results will be discusses further in a later section.

A



B



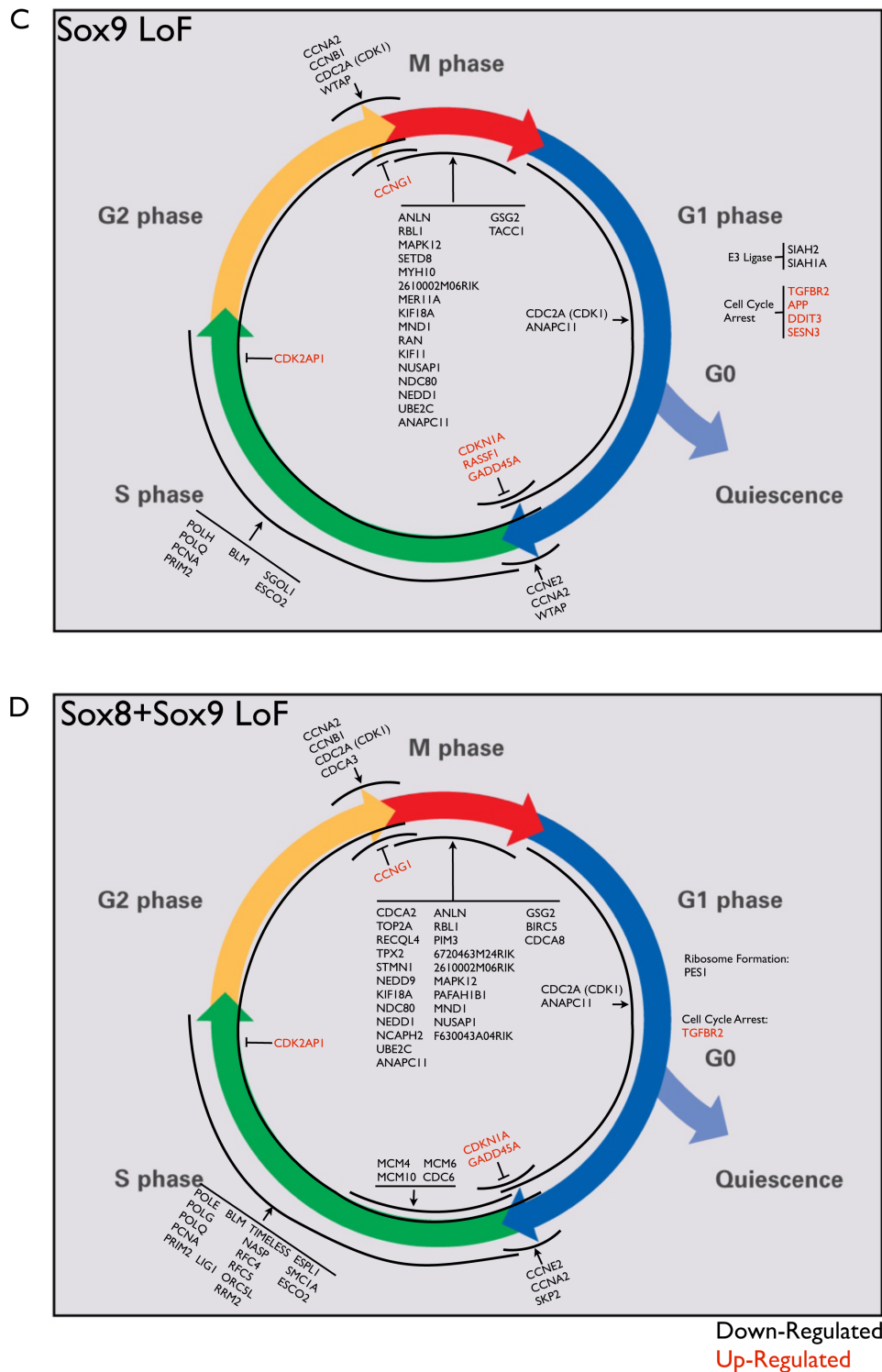


Figure 4.7: SoxE downstream genes are involved in cell cycle regulation

List of SoxE downstream genes involved in cell cycle regulation based on functional annotation analysis (A). The distribution of Sox8 (B), Sox9 (C) and SoxE (D) downstream cell cycle related genes in the cell cycle process.

4.3.6 Sox9 directly regulate cell cycle related genes

Taking the advantage that the ChIP-Seq experiment using NS5 cells as starting material was performed previously (Ben Martynoga, unpublished data), I identified 513 direct targets of Sox9. Furthermore, 21 out of 41 cell cycle related genes, which was identified by the expression microarray were direct targets of Sox9 (Figure 4.8 A). This confirmed that Sox9 was directly involved in NSC cell cycle regulation *in vitro*. Furthermore, Sox9 targets involved in cell proliferation were involved in multiple stages of the cell cycle process (Figure 4.8 B).

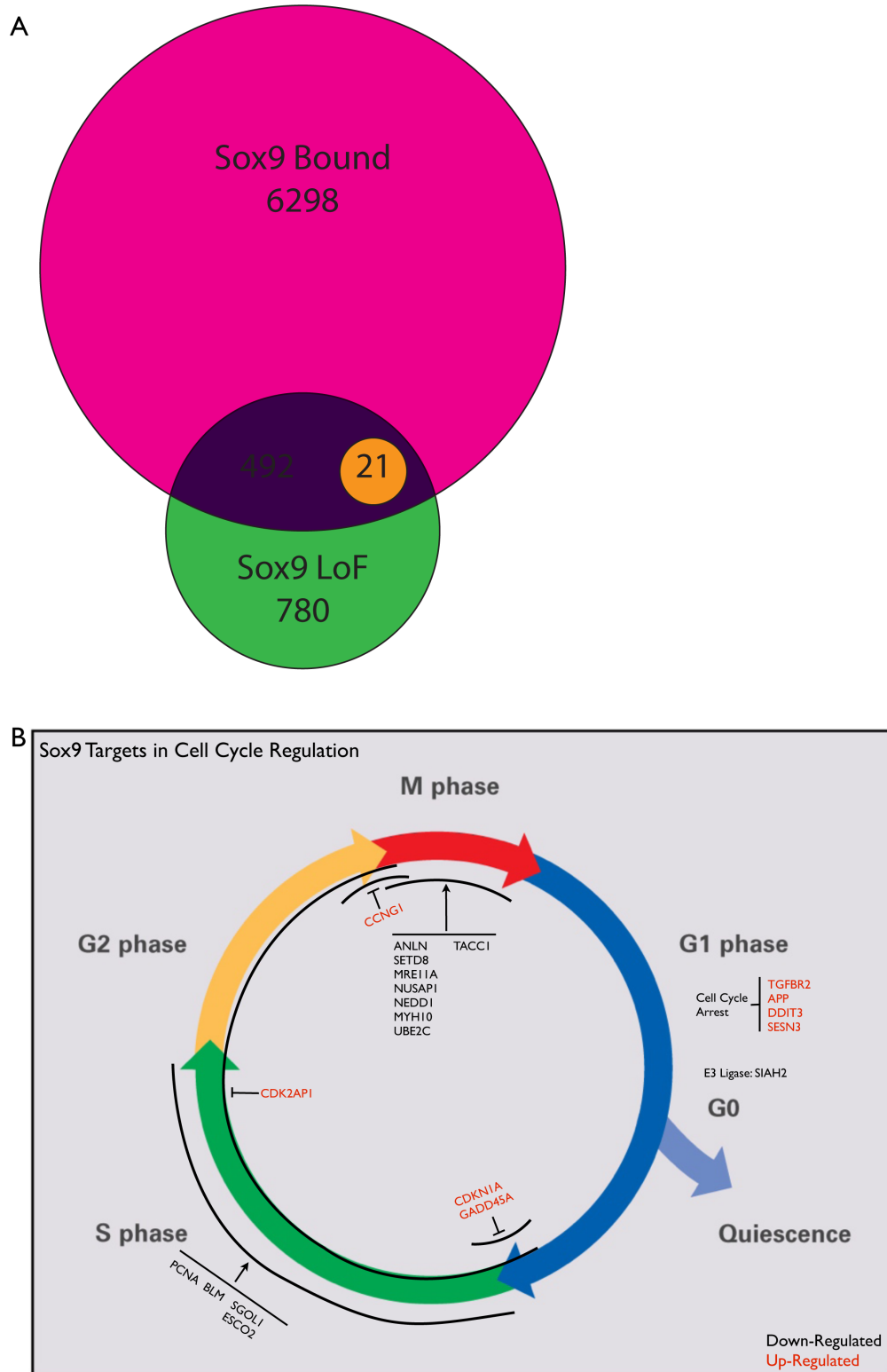


Figure 4.8: Sox9 target genes were involved in cell cycle regulation
Identification of Sox9 target genes (A). The distribution of the 21 cell proliferation related Sox9 target genes in cell cycle process (B).

4.4 Uncovering molecular mechanisms of SoxE in regulating NSC proliferation

My pervious experiments have shown SoxE activity is essential to maintain NSC proliferation *in vitro* (Figure 4.4). Genome-wide gene expression microarray and ChIP-Seq analysis are consistent with this observation and further suggest that SoxE regulates cell cycle related genes in a direct manner (at least for Sox9, Figure 4.7 and Figure 4.8). However, the molecular mechanisms of SoxE in regulating cell proliferation remain unclear and became my next question. However, downstream and target genes of Sox8 and Sox9 related with cell proliferation are distributed across the entire cell cycle process (Figure 4.7 and Figure 4.8). In order to identify the most critical downstream or target genes involved in this process, subsequently the molecular mechanism, the cell cycle dynamic needed to be assessed. Immunocytochemistry and flow cytometry were used to examine the cell cycle dynamic under the SoxE LoF conditions.

4.4.1 SoxE is involved in G1/S transition

To assess SoxE activity in controlling cell cycle dynamics, NS5 cells were transfected with shRNA against Sox8 or/and Sox9. EdU containing medium was applied 48h post transfection and cells were fixed at different time points (Figure 4.9 A).

Under the control condition, the labelling index (LI) (the number of GFP-EdU double positive cells divided by the total number of GFP-positive cells) progressively increased up to 13.38h (point A, Figure 4.9 B). This corresponded to the growth fraction (GF), indicating that 88.57% of the GFP-positive cells were proliferating. The data (Figure 4.9 B, blue line) could be described by two regression lines: the linear line increased in cell labeling between 20 min and 13.38h ($y=4.754 \cdot x + 25.06$) and the horizontal line between 13.38h and 30h. The time (point A) when the two lines intersect equals $T_c - T_s = 13.38h$. The intersect of the linear line with the y axis was defined as the initial labeling index (LI0) and corresponds to $GF \times T_s / T_c$. For my data, LI0 was 25.06, suggesting 25.06% of the GFP-positive cells were in the S-phase at T0. From these three equations, I calculated a T_c of 18.77h and a T_s of 5.28h for the GFP-positive cells (Figure 4.9

B). Under the SoxE knock down condition, the LI slowly increased but did not reach the GF within 30h. The slow increasing linear lines (Sox8 KD: $y=0.4330*x + 7.272$; Sox9 KD: $y=0.8159*x + 10.30$; Sox8-Sox9 double KD: $y=0.2072*x + 6.057$) indicated that the LI0 for Sox8 KD, Sox9 KD and Sox8-Sox9 double KD were 7.272, 10.30 and 6.057. The GF for cells under LoF conditions was estimated as 88.58% and the Tc and Ts were calculated as described above (Figure 4.9 B).

These data suggest that SoxE is required for G1-phase progression or/and G1/S transition during NSC proliferation. This will be discussed further in a later section.

4.4.2 SoxE activity is not required for mitosis

To assess whether SoxE activity was involved in mitosis, antibody against pHH3 (a marker for late G2 and M-phase) was used to visualize the process.

Immunocytochemistry was performed 48h post-transfection (Figure 4.10 A). Under the control condition, pHH3 expression was detected at late G2/early M phase (prophase, Figure 4.10 B) and late M phase (cytokinesis, Figure 4.10 B'). SoxE LoF did not have clear effect on the M-phase as both prophase and telophase/cytokinesis appeared to be normal under single and double SoxE KD conditions (Figure 4.10 C to E'). However, SoxE LoF reduced the percentage of cells entered the M-phase. Under Sox8 single and Sox8-Sox9 double KD conditions, the percentage was reduced by more than 60%. The percentage of cells entered the M-phase was reduced by 50% under the Sox9 LoF conditions (Figure 4.10 F).

The data suggest that SoxE is not involved in mitosis directly and this will be further discussed later.

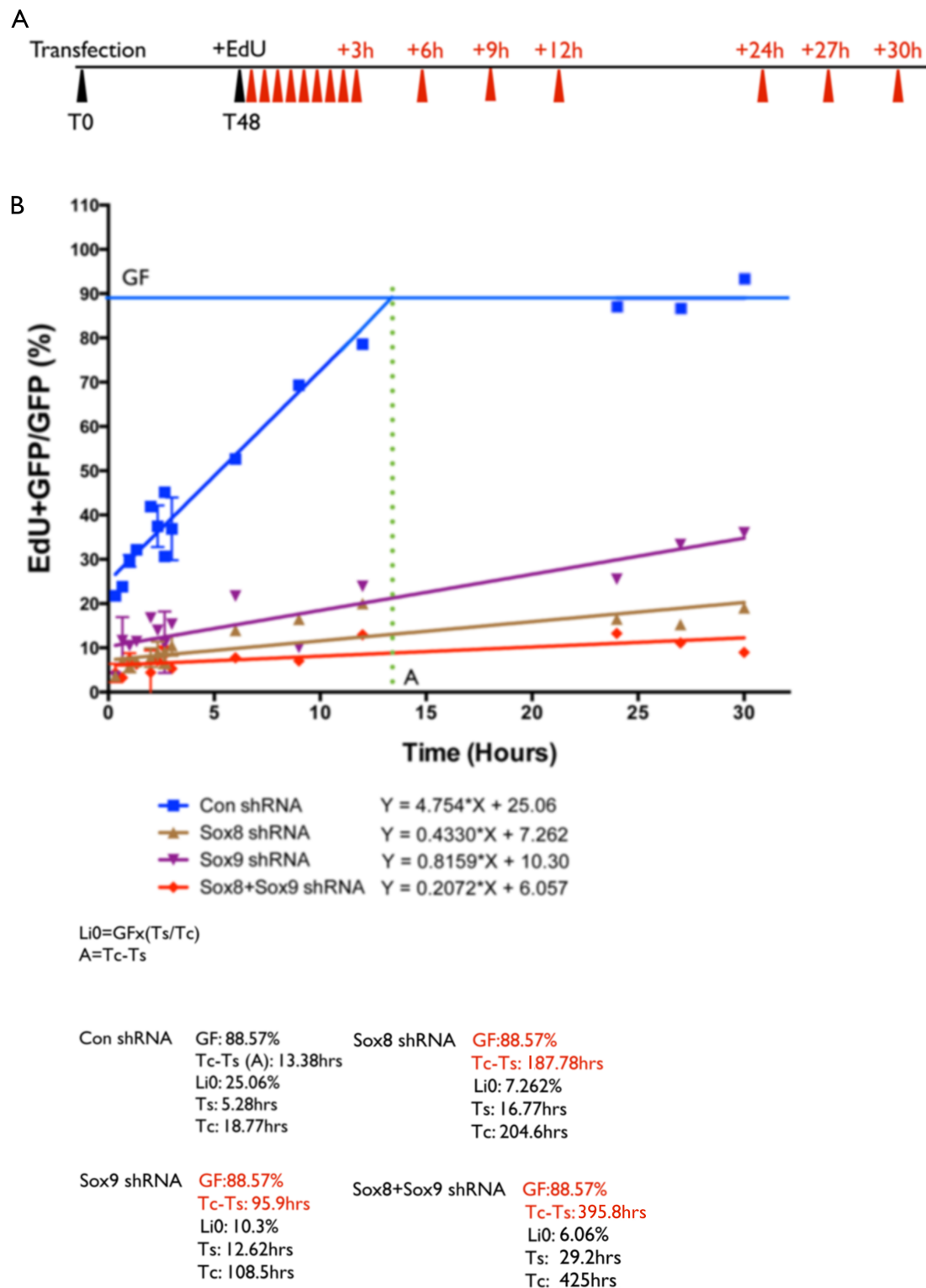


Figure 4.9: SoxE is required for G1/S transition during cell cycle progression

Experimental design of the cumulative labelling (CL) method (A). Length of cell cycle phases under control (blue line), Sox8 KD (brown line), Sox9 KD (purple line) and Sox8-Sox9 double KD (red line) conditions were calculated using the LI data.

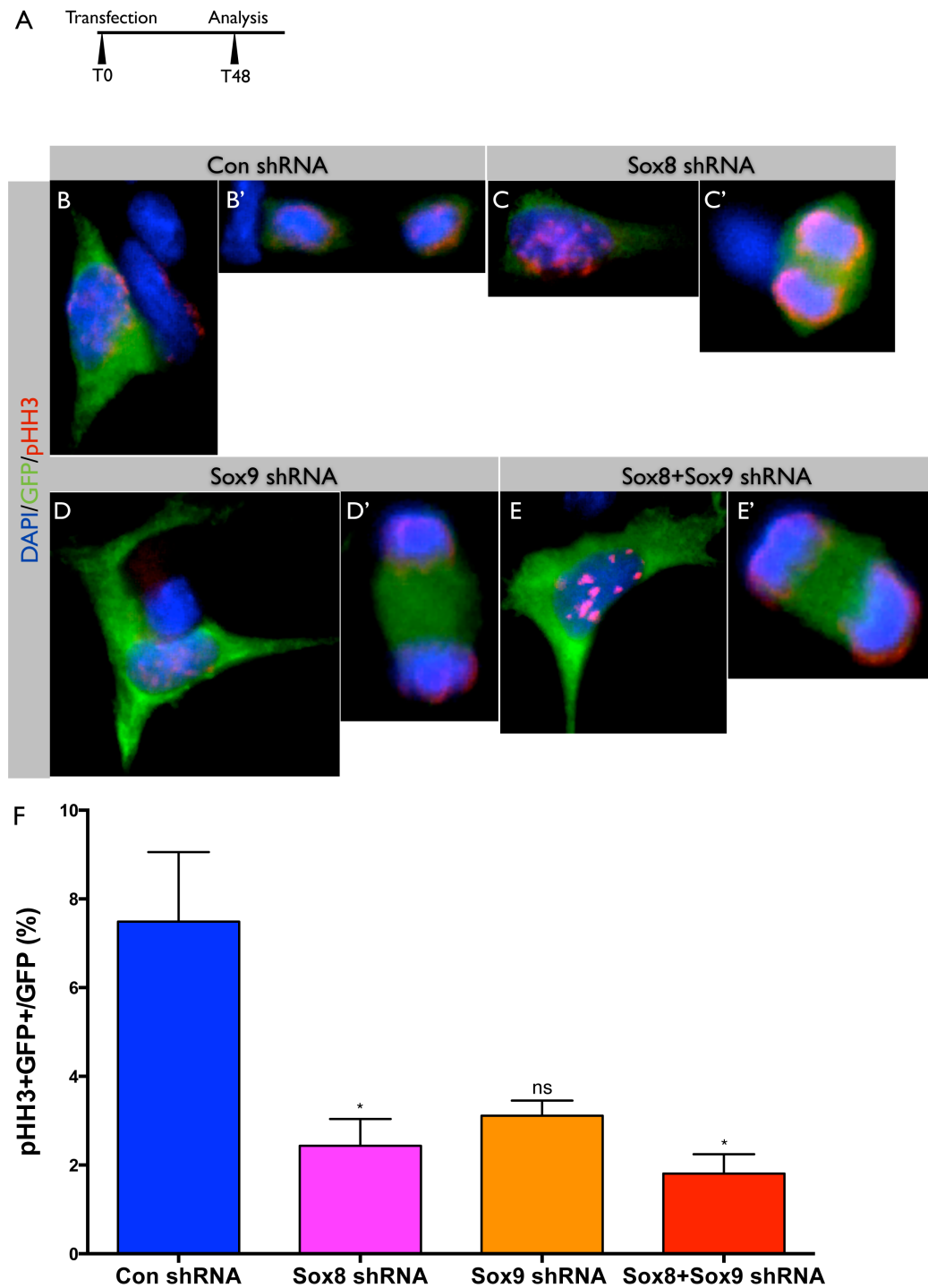


Figure 4.10: SoxE activity was not involved in mitosis

Experimental design (A). Expression of pHH3 (red) in late G2/early M phase and late M phase under control (B and B'), Sox8 KD (C and C'), Sox9 KD (D and D') and Sox8-Sox9 double KD (E and E'). The quantification graph shows percentage pHH3-GFP double positive cells 48h post-transfection (F). Data shown are mean of three independent biological replicates. Error bar: mean \pm SEM. t test: * $p < 0.05$.

4.4.3 Cell cycle analysis using fluorescence-activated cell sorting (FACS) analysis

In parallel with immunocytochemistry, FACS analysis was used as a second approach to assess cell cycle dynamic. Based on DNA content, FACS analysis is able to define the percentage of cells at G1/G0, S and G2/M phases within the population. The accumulation of cells in a particular phase of cell cycle after SoxE KD would be helpful to identify potential candidates for understanding molecular mechanism of SoxE in regulating NSC proliferation. My preliminary data suggested the cells were potentially stuck at G1/G0-phase and I was expecting an accumulation of G1/G0 cells during the experiment (see 4.3.2 for details).

NS5 cells transfected with Sox8 or/and Sox9 shRNA were collected 72h post transfection with 56h puromycine selection (Figure 4.11 A). FACS analysis profile showed the distribution of transfected cells in different phases during cell cycle under the control condition (G0/G1: 73%; S: 12.7%; G2/M: 12%, Figure 4.11 B). Surprisingly, the distribution profiles for both single and double knock down conditions were similar to control (Figure 4.11 B', B'' and B'''). No clear accumulation of cells in any of the cell cycle phases was observed under the KD conditions (Figure 4.11 C).

The data from FACS analysis was not consistent with earlier experiments, which showed a smaller number of cells had entered S and M phases (Figure 4.9, see 4.3.2 for details). Till this point, I do not have other explanations but technique issues (eg: cell fixation) to explain the results.

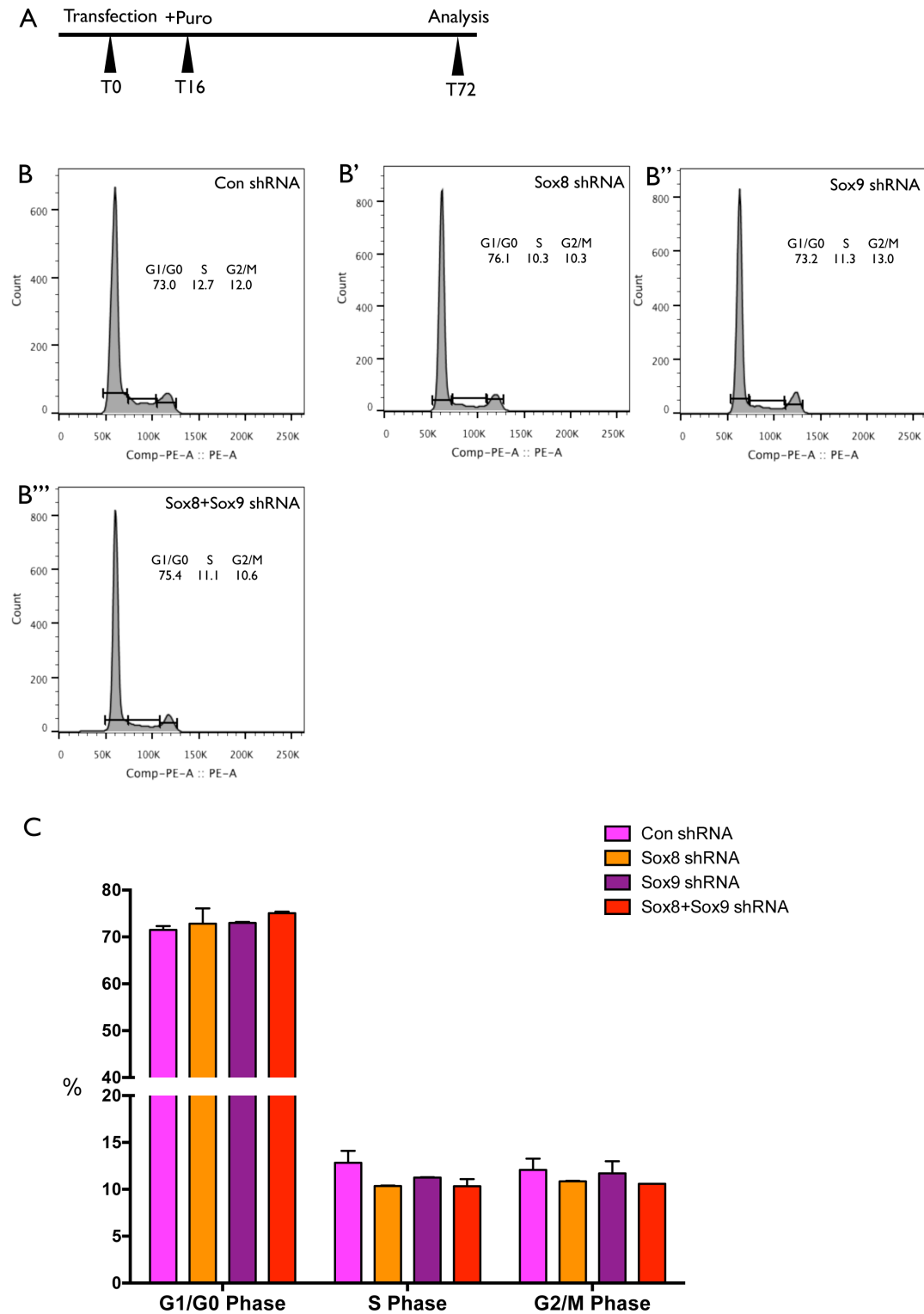


Figure 4.11: Cell cycle analysis under SoxE KD conditions

Experimental design (A). FACS analysis profiles for control (B), Sox8 KD (B'), Sox9 KD (B'') and Sox8-Sox9 double KD (B'''). The quantification graph shows the distribution of transfected cells in different phases of the cell cycle. Data shown are mean of two independent biological replicates. Error bar: mean \pm SEM.

4.5 Regulation between SoxE and Ascl1 and within the SoxE members

As mentioned earlier, Sox8 regulates Ascl1 activity in an enhancer dependent manner (Cristina Minieri, unpublished data) and interact with Ascl1 at the protein level (Debbie van den Berg, unpublished data). q-PCR was used to assess whether Sox8 and Sox9 were able to regulate Ascl1 expression directly. I also noticed cross regulation between the two SoxE members.

RNA samples were collected 48h and 72h post-transfection with 32h and 56h puromycine selection (Figure 4.12 A). q-PCR analysis showed a significant reduction of Ascl1 expression under both single and double KD conditions at both time points (for Sox8 KD at 48h, a 25% reduction was observed). This suggested the possibility SoxE positively regulate Ascl1 expression at least at the RNA level (Figure 4.12 B).

Furthermore, under the Sox8 KD condition, the expression of Sox9 was up-regulated by 1.4-fold 48h post-transfection. At the 72h time point, a 3.5-fold increase was observed. This suggested Sox8 repressed Sox9 expression at the RNA level (Figure 4.12 C). On the other hand, Sox8 expression was not effected by Sox9 DK (data not shown). These results will be discussed in a later section.

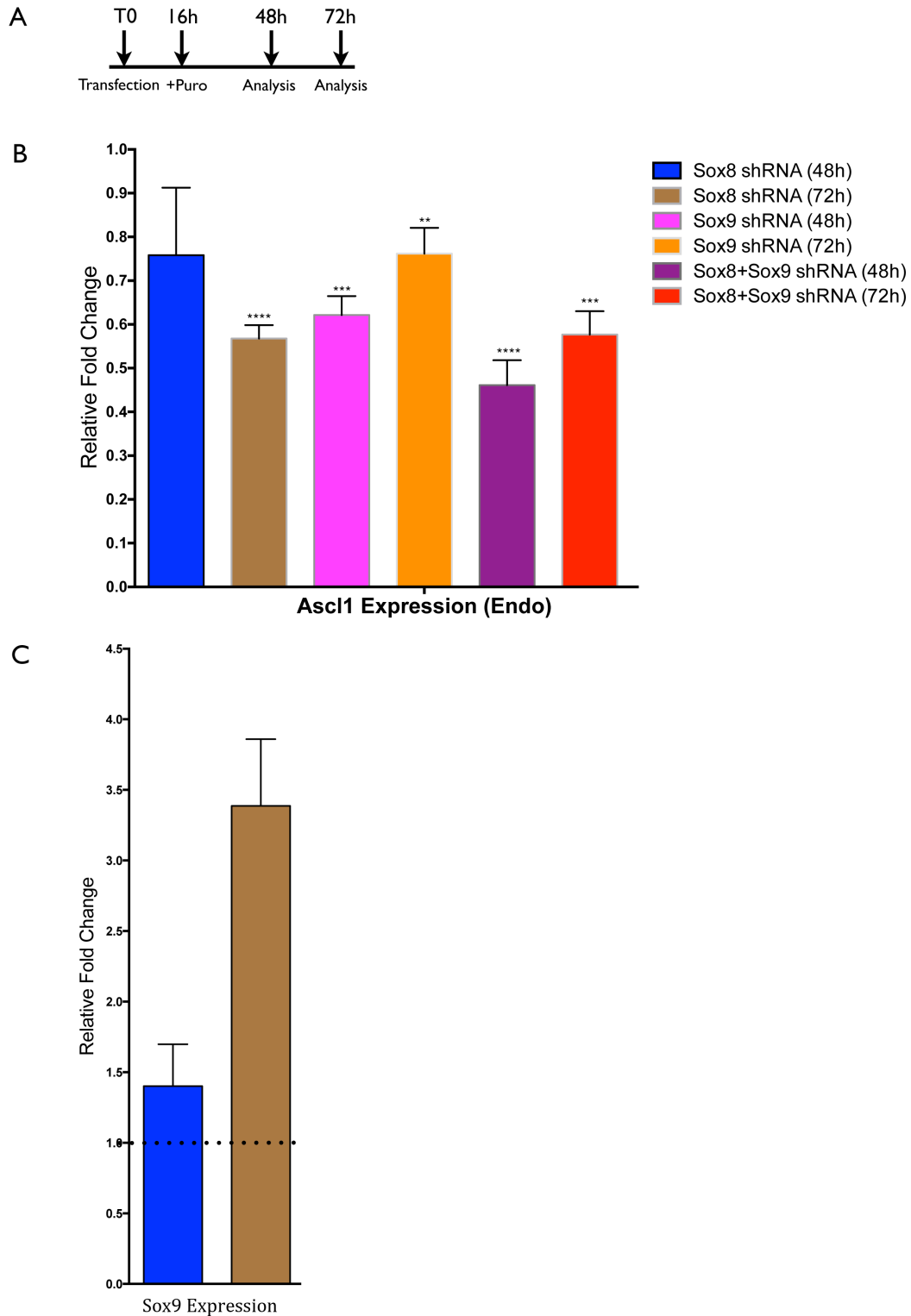


Figure 4.12: Regulation between SoxE members and Ascl1

Experimental design (A). q-PCR analysis to assess *Ascl1* expression 48h and 72h post-transfection under SoxE KD conditions (B). q-PCR analysis to assess *Sox9* expression under the *Sox8* LoF condition. Data shown are mean of three independent biological replicates. Error bar: mean \pm SEM. t test: ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$

4.6 Discussion

In this chapter, I have shown that ectopic expression of Sox8 or Sox9 inhibit Ascl1-induced neuronal differentiation *in vitro*. Furthermore, my preliminary data has also shown that Sox8 and Sox9 have essential functions in maintaining NSC proliferation *in vitro*, potentially via regulating G1 phase progression or/and G1/S transition. Here, I will discuss these preliminary findings.

4.6.1 SoxE inhibits Ascl1-mediated NSC differentiation *in vitro*

My GoF studies have shown that Sox8 and Sox9 negatively regulated Ascl1-induced NSC neuronal differentiation *in vitro*. This observation is consistent with a number of published studies, which indicate Sox8 or/and Sox9 inhibit cell differentiation in different tissues. In CNS, Sox9 GoF inhibits neuronal differentiation in developing mouse telencephalon (Scott et al, 2010). In other tissues, Sox8 acts as negatively regulator osteoblast differentiation (Schmidt et al, 2005); both Sox8 and Sox9 inhibit MyoD-induced myogenesis (Schmidt et al, 2003). Sox8 or Sox9 GoF reduces the Ascl1-induced NSC neuronal differentiation *in vitro* suggests two possibilities: SoxE down regulate Ascl1 expression directly or inhibits Ascl1-induced activation of neuronal specific targets.

Genome wide expression microarray data has shown that Sox8 is up regulated during Ascl1-induced NS5 cell differentiation (Ben Martynoga, unpublished data). Furthermore, SoxE LoF experiment shows a reduction of Ascl1 expression (Figure 4.12 B). These suggest that Sox8 and Ascl1 positively regulate each other at the transcriptional level. I do not have direct evidence to indicate that Sox9 is regulated by Ascl1 expression. Indeed, Sox9 is down regulated during Ascl1-induced NS5 cell differentiation (Ben Martynoga, unpublished data), but this potentially due to the up-regulation of Sox8. On the other hand, luciferase reporter assay indicates that Sox8 GoF inhibits Ascl1 activity in an enhancer dependent manner (Cristina Minieri, unpublished data). These data suggest the possibility that Sox8 inhibits Ascl1-induced neuronal differentiation via blocking activities of Ascl1 target genes *in vitro*. No direct evidence to suggest Sox9 is able to control the activity of Ascl1 targets.

Here, the data raise the possibility that SoxE (at least Sox8) inhibit Ascl1-induced neuronal differentiation via direct inhibition of Ascl1 expression or/and controlling Ascl1 target genes.

4.6.2 Sox8 and Sox9 are essential and have partial functional redundancy in maintaining NSC proliferation *in vitro*

My preliminary data has shown that Sox8 and Sox9 have critical functions to maintain NSC in a proliferative state. The significant reduction of cell proliferation (Figure 4.4 and Figure 4.9) under single KD conditions indicates that Sox8 and Sox9 cannot compensate each other to maintain NSC proliferation. On the other hand, the double KD indeed has an enhanced phenotype in inhibiting NSC proliferation comparing with Sox8 or Sox9 single knock down (Figure 4.4 and Figure 4.9). These data indicate that SoxE members only have partial functional redundancy in maintaining NSC proliferation. This is also indicated in my genome wide expression microarray data (see 4.6.3 for further discussion).

The observation that Sox8 and Sox9 have partial functional redundancy raises two hypotheses to predict the molecular mechanisms of Sox8 and Sox9 in regulating cell proliferation. Firstly, Sox8 and Sox9 share common/similar downstream genes, which are involved in maintaining NSC proliferation (eg: Cdc2a and Cdkn1a; Figure 4.7). The enhanced phenotype in inhibiting proliferation in the double knock down may due to further reduction of SoxE activities. Secondly, Sox8 and Sox9 regulate different sets of genes that give similar phenotype. For example, Skp2, an E3 ubiquitin ligase that is essential for G1/S transition is down regulated under Sox8 LoF condition (Figure 4.7). Gadd45a, which inhibits cells entering S-phase is up regulated under Sox9 LoF condition (Figure 4.7). The enhanced phenotype in the double knock down may due to the combination of the two activities. However, I have not had any direct evidence to support either hypothesis.

4.6.3 Identification and functional annotation of SoxE downstream genes

Genome wide expression microarrays was used to detect gene expression changes under SoxE LoF conditions. Under the single KD conditions, genes

deregulated by either Sox8 LoF or Sox9 LoF indicates that the two members regulate different set of genes and subsequently have different functions. Sox8 and Sox9 have highly conserved DNA binding HMGs and binding motif (Figure 4.1), the difference in optimal binding sequence partly explain the observation from the microarray data (Figure 4.5) and the GO analysis (Figure 4.6).

Genes deregulated under the double KD only are redundantly regulated by Sox8 and Sox9. However, 714 genes are deregulated under the single KD but not the double KD (Figure 4.5). There are two potential sources for this set of genes. Firstly, genes are deregulated to opposite directions under Sox8 and Sox9 single KD conditions. For example, Sox9 is up regulated under the Sox8 KD condition. Only 19 genes are regulated in this manner. Variations in KD efficiency between biological replicates and between single and double KD (Figure 4.5) is major source for this set of genes. This technique issue cannot be eliminated under current *in vitro* KD system. Similar observation can be found for cell cycle genes (Figure 4.7).

4.6.4 SoxE activity is involved in G1-phase progression and G1/S transition, but not mitosis

The CL experiment has shown that cells under SoxE KD conditions (both singles and double KD) have lower LI0 values and slower LI increase, which unable to reach GF of the control population within the given time. These suggest that the percentage of cells has entered S-phase is much lower under SoxE KD conditions and the cells are unable/slowly pass G1/S checkpoint to complete cell cycle (Figure 4.9). Furthermore, as cells are able to go through mitosis normally under SoxE KD conditions, suggesting Sox8 and Sox9 are not directly involved in regulating M-phase (Figure 4.10).

The data raise the possibility that SoxE members have critical functions in regulating G1 progression or/and G1/S transition during the cell cycle. Pervious studies have shown that Sox9 regulates the proliferation of primary cells (MEFs) via controlling cells cycle progression from G1 to S-phase (Matheu et al, 2012). My data is consistent with the pervious observation. Although the experiments were performed using two different *in vitro* systems, Sox9 may still potentially regulate similar target genes to control cell cycle progression in both systems. In

combination with my expression microarray data, potential candidate(s) for uncovering the molecular mechanism should be primarily selected from genes involved in G1-phase progression and G1/S transitions (eg: Cdk1, Cdkn1a, Skp2; Figure 4.7). However, till this point, whether similar/same molecular mechanism(s) is used by Sox8 and Sox9 to control cell cycle progression is unclear and further investigation is required. However, whether the cells entered the S-phase are able to complete the cell cycle has yet to be confirmed using double thymidine analog (DA) method.

Furthermore, the lower percentage of cells in S phase (lower LI0 value) under the KD conditions suggests that a sub population of cells is unable to proliferate and stuck at G1/G0 phase (as SoxE is potentially control G1 progression or/and G1/S transition. However, whether the cells exit the cell cycle and enter G0 phase is unclear). On the other hand, the slow increase of LI under the KD conditions indicates a sub population of transfected cells indeed can proliferate, but slowly (Figure 4.9). These suggest that the transfected cells can be divided into two populations: the non-proliferating and slow proliferating populations. This can be partly explained by the experimental approach. One of disadvantages of the shRNA-mediated KD is the variation in KD efficiency. Even within a single transfection reaction, the KD efficiency between individual transfected cells is potentially different, which is indicated by the variation in the GFP signal intensity (Figure 4.4). The cells with high KD efficiency are potentially the source of the non-proliferating cells and cells with low KD efficiency may still proliferate, but slowly.

My preliminary results have shown that SoxE is not directly required for mitosis, as both prophase and telophase/cytokinesis are not affected under SoxE KD conditions (Figure 4.10). This indicates that SoxE downstream/target genes related with M-phase progression are not directly responsible to control NSC proliferation (Figure 4.7). Indeed, the percentage of transfected cells entering S-phase is reduced by more than 50% (Figure 4.9), which can be partly explained by the presence of non-proliferating cells (see above for details). Furthermore, the expression of pHH3 starts from late G2 phase and the reduced number of pHH3-positive cells suggests less cells have entered (late) G2 phase. This is

consistent with the hypothesis that SoxE regulate NSC cell cycle via controlling G1 phase progression or/and G1/S transition.

4.6.5 SoxE and Ascl1 in regulating NSC proliferation

q-PCR analysis has shown that Sox8 KD increase Sox9 expression at the transcriptional level. This suggests Sox8 potentially acts upstream and represses Sox9 expression. Previous studies have shown that Sox8 knock out (KO) transgenic mice do not have any developmental defects at both embryonic and adult stages, despite Sox8 is strongly expression in many tissues, including CNS. Lack of developmental defects in Sox8 deficient mice may due to functional redundancy between Sox8 and Sox9. The only noticeable effect of the Sox8 deficient animal is the weight loss, which is potentially an indication of reduced cell proliferation.

Earlier studies have shown that Ascl1 is essential to maintain NSC proliferation by regulating cell cycle genes in a direct manner (Castro et al, 2011). q-PCR analysis has shown that Sox8 and Sox9 positively regulate Ascl1 expression, which suggests SoxE is required to maintain Ascl1 expression and subsequently the proliferation of NSCs. Furthermore, a number of Ascl1 target genes involved in cell proliferation act downstream of SoxE members. For example, Cdk1, which is involved in G1 phase progression, is down regulated under Sox8 and Sox9 LoF condition. Skp2, an E3 ubiquitin ligase involved in G1/S transition, is down regulated under the Sox8 KD condition. Both Cdk1 and Skp2 are direct targets of Ascl1. These data suggest the possibility that SoxE not only regulate Ascl1 expression directly, also the expression of Ascl1 targets to maintain NSC proliferation.

Conclusions

As a proneural transcriptional factor, Ascl1 plays essential roles in activating neuronal specific programs in both direct reprogramming of MEFs and NSC differentiation. Here, our GoF and LoF studies indicated that Sox8 was involved in Ascl1-mediated MEFs reprogramming and the expression of Sox8 during the process was transient. Furthermore, our data suggested that Sox8 and Sox9 were essential to maintain NSC in a proliferating state *in vitro*. Genome wide expression profiling had made the first step to understand mechanisms, which underlie the regulatory roles of SoxE in NSC proliferation. Preliminary studies also suggested the possibility that SoxE could regulate Ascl1 activity, which in turn could have impacts on both MEFs reprogramming and NSC proliferation.

Bibliography

Ambasudhan, R. *et al.* Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions. *Cell stem cell* **9**, 113-118 (2011).

Anderson, D.J. Cellular and molecular biology of neural crest cell lineage determination. *Trends in genetics : TIG* **13**, 276-280 (1997).

Artavanis-Tsakonas, S., Matsuno, K. & Fortini, M.E. Notch signaling. *Science* **268**, 225-232 (1995).

Bergsland, M., Werme, M., Malewicz, M., Perlmann, T. & Muhr, J. The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes & development* **20**, 3475-3486 (2006).

Berninger, B., Guillemot, F. & Gotz, M. Directing neurotransmitter identity of neurones derived from expanded adult neural stem cells. *The European journal of neuroscience* **25**, 2581-2590 (2007).

Bertrand, N., Castro, D.S. & Guillemot, F. Proneural genes and the specification of neural cell types. *Nature reviews. Neuroscience* **3**, 517-530 (2002).

Bondurand, N. *et al.* Human Connexin 32, a gap junction protein altered in the X-linked form of Charcot-Marie-Tooth disease, is directly regulated by the transcription factor SOX10. *Human molecular genetics* **10**, 2783-2795 (2001).

Caiazzo, M. *et al.* Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* **476**, 224-227 (2011).

Campuzano, S. & Modolell, J. Patterning of the Drosophila nervous system: the achaete-scute gene complex. *Trends in genetics : TIG* **8**, 202-208 (1992).

Casarosa, S., Fode, C. & Guillemot, F. Mash1 regulates neurogenesis in the ventral telencephalon. *Development* **126**, 525-534 (1999).

Castro, D.S. *et al.* A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. *Genes & development* **25**, 930-945 (2011).

Chen, H.Z., Tsai, S.Y. & Leone, G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nature reviews. Cancer* **9**, 785-797 (2009).

Cheung, M. & Briscoe, J. Neural crest development is regulated by the transcription factor Sox9. *Development* **130**, 5681-5693 (2003).

Chin, M.H. *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell stem cell* **5**, 111-123 (2009).

Costa, R.H. FoxM1 dances with mitosis. *Nat Cell Biol* **7**, 108-110 (2005).

Cowan, C.A., Atienza, J., Melton, D.A. & Eggan, K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* **309**, 1369-1373 (2005).

Davis, R.L. & Turner, D.L. Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning. *Oncogene* **20**, 8342-8357 (2001).

Duncan, M., DiCicco-Bloom, E.M., Xiang, X., Benezra, R. & Chada, K. The gene for the helix-loop-helix protein, Id, is specifically expressed in neural precursors. *Developmental biology* **154**, 1-10 (1992).

Feng, B. *et al.* Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* **11**, 197-203 (2009).

Fode, C. *et al.* A role for neural determination genes in specifying the dorsoventral identity of telencephalic neurons. *Genes & development* **14**, 67-80 (2000).

Garcia-Bellido, A. Genetic Analysis of the Achaete-Scute System of DROSOPHILA MELANOGASTER. *Genetics* **91**, 491-520 (1979).

Ghislain, J. & Charnay, P. Control of myelination in Schwann cells: a Krox20 cis-regulatory element integrates Oct6, Brn2 and Sox10 activities. *EMBO reports* **7**, 52-58 (2006).

Gonzalez, F., Romani, S., Cubas, P., Modolell, J. & Campuzano, S. Molecular analysis of the asense gene, a member of the achaete-scute complex of Drosophila melanogaster, and its novel role in optic lobe development. *The EMBO journal* **8**, 3553-3562 (1989).

Gopinathan, L., Ratnacaram, C.K. & Kaldis, P. Established and novel Cdk/cyclin complexes regulating the cell cycle and development. *Results and problems in cell differentiation* **53**, 365-389 (2011).

Goshima, Y., Nakamura, F., Strittmatter, P. & Strittmatter, S.M. Collapsin-induced growth cone collapse mediated by an intracellular protein related to UNC-33. *Nature* **376**, 509-514 (1995).

Goulding, S.E., White, N.M. & Jarman, A.P. cato encodes a basic helix-loop-helix transcription factor implicated in the correct differentiation of Drosophila sense organs. *Developmental biology* **221**, 120-131 (2000).

Goulding, S.E., zur Lage, P. & Jarman, A.P. amos, a proneural gene for Drosophila olfactory sense organs that is regulated by lozenge. *Neuron* **25**, 69-78 (2000).

Grove, E.A., Tole, S., Limon, J., Yip, L. & Ragsdale, C.W. The hem of the embryonic cerebral cortex is defined by the expression of multiple Wnt genes and is compromised in Gli3-deficient mice. *Development* **125**, 2315-2325 (1998).

Gubbay, J. *et al.* A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346**, 245-250 (1990).

Guillemot, F. Cellular and molecular control of neurogenesis in the mammalian telencephalon. *Current opinion in cell biology* **17**, 639-647 (2005).

Guillemot, F. *et al.* Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* **75**, 463-476 (1993).

Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595-601 (2009).

Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature biotechnology* **26**, 795-797 (2008).

Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature biotechnology* **26**, 1269-1275 (2008).

Jessell, T.M. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nature reviews. Genetics* **1**, 20-29 (2000).

Jin, B., Li, Y. & Robertson, K.D. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes & cancer* **2**, 607-617 (2011).

Kamachi, Y., Uchikawa, M. & Kondoh, H. Pairing SOX off: with partners in the regulation of embryonic development. *Trends in genetics : TIG* **16**, 182-187 (2000).

Kim, J., Lo, L., Dormand, E. & Anderson, D.J. SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells. *Neuron* **38**, 17-31 (2003).

Kwan, K.Y., Sestan, N. & Anton, E.S. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* **139**, 1535-1546 (2012).

LeBlanc, S.E., Ward, R.M. & Svaren, J. Neuropathy-associated Egr2 mutants disrupt cooperative activation of myelin protein zero by Egr2 and Sox10. *Molecular and cellular biology* **27**, 3521-3529 (2007).

Li, B., Carey, M. & Workman, J.L. The role of chromatin during transcription. *Cell* **128**, 707-719 (2007).

Lo, L., Dormand, E., Greenwood, A. & Anderson, D.J. Comparison of the generic neuronal differentiation and neuron subtype specification functions of mammalian achaete-scute and atonal homologs in cultured neural progenitor cells. *Development* **129**, 1553-1567 (2002).

Marson, A. *et al.* Wnt signaling promotes reprogramming of somatic cells to pluripotency. *Cell stem cell* **3**, 132-135 (2008).

Massari, M.E. & Murre, C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Molecular and cellular biology* **20**, 429-440 (2000).

Mattar, P. *et al.* Basic helix-loop-helix transcription factors cooperate to specify a cortical projection neuron identity. *Molecular and cellular biology* **28**, 1456-1469 (2008).

Mertin, S., McDowall, S.G. & Harley, V.R. The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic acids research* **27**, 1359-1364 (1999).

Mikkelsen, T.S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49-55 (2008).

Mizuguchi, R. *et al.* Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron* **31**, 757-771 (2001).

Murre, C., McCaw, P.S. & Baltimore, D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**, 777-783 (1989).

Murre, C. *et al.* Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-544 (1989).

O'Donnell, M., Hong, C.S., Huang, X., Delnicki, R.J. & Saint-Jeannet, J.P. Functional analysis of Sox8 during neural crest development in *Xenopus*. *Development* **133**, 3817-3826 (2006).

Pang, Z.P. *et al.* Induction of human neuronal cells by defined transcription factors. *Nature* **476**, 220-223 (2011).

Parras, C.M. *et al.* Divergent functions of the proneural genes Mash1 and Ngn2 in the specification of neuronal subtype identity. *Genes & development* **16**, 324-338 (2002).

Perez, S.E., Rebelo, S. & Anderson, D.J. Early specification of sensory neuron fate revealed by expression and function of neurogenins in the chick embryo. *Development* **126**, 1715-1728 (1999).

Pevny, L.H., Sockanathan, S., Placzek, M. & Lovell-Badge, R. A role for SOX1 in neural determination. *Development* **125**, 1967-1978 (1998).

Pfisterer, U. *et al.* Direct conversion of human fibroblasts to dopaminergic neurons. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 10343-10348 (2011).

Pierfelice, T., Alberi, L. & Gaiano, N. Notch in the vertebrate nervous system: an old dog with new tricks. *Neuron* **69**, 840-855 (2011).

Rallu, M. *et al.* Dorsoventral patterning is established in the telencephalon of mutants lacking both Gli3 and Hedgehog signaling. *Development* **129**, 4963-4974 (2002).

Ross, S.E., Greenberg, M.E. & Stiles, C.D. Basic helix-loop-helix factors in cortical development. *Neuron* **39**, 13-25 (2003).

Rossant, J. Stem cells and early lineage development. *Cell* **132**, 527-531 (2008).

Ruiz, S. *et al.* A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. *Current biology : CB* **21**, 45-52 (2011).

Sandberg, M., Kallstrom, M. & Muhr, J. Sox21 promotes the progression of vertebrate neurogenesis. *Nature neuroscience* **8**, 995-1001 (2005).

Sarkar, A. & Hochedlinger, K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell stem cell* **12**, 15-30 (2013).

Schmidt, K., Glaser, G., Wernig, A., Wegner, M. & Rosorius, O. Sox8 is a specific marker for muscle satellite cells and inhibits myogenesis. *The Journal of biological chemistry* **278**, 29769-29775 (2003).

Schmidt, K. *et al.* The high mobility group transcription factor Sox8 is a negative regulator of osteoblast differentiation. *The Journal of cell biology* **168**, 899-910 (2005).

Schuurmans, C. & Guillemot, F. Molecular mechanisms underlying cell fate specification in the developing telencephalon. *Current opinion in neurobiology* **12**, 26-34 (2002).

Scott, C.E. *et al.* SOX9 induces and maintains neural stem cells. *Nature neuroscience* **13**, 1181-1189 (2010).

Sherr, C.J. & Roberts, J.M. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes & development* **13**, 1501-1512 (1999).

Shi, Y. *et al.* Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell stem cell* **3**, 568-574 (2008).

Shi, Y. *et al.* A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell stem cell* **2**, 525-528 (2008).

Silva, J. *et al.* Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS biology* **6**, e253 (2008).

Sinclair, A.H. *et al.* A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346**, 240-244 (1990).

Son, E.Y. *et al.* Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell stem cell* **9**, 205-218 (2011).

Stolt, C.C., Lommes, P., Friedrich, R.P. & Wegner, M. Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy. *Development* **131**, 2349-2358 (2004).

Stolt, C.C. *et al.* The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes & development* **17**, 1677-1689 (2003).

Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes & development* **12**, 599-606 (1998).

Suzuki, M.M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nature reviews. Genetics* **9**, 465-476 (2008).

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada, T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Current biology : CB* **11**, 1553-1558 (2001).

Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).

Taylor, S.M. & Jones, P.A. Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell* **17**, 771-779 (1979).

Toresson, H., Potter, S.S. & Campbell, K. Genetic control of dorsal-ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2. *Development* **127**, 4361-4371 (2000).

Uwanogho, D. *et al.* Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes suggests an interactive role in neuronal development. *Mechanisms of development* **49**, 23-36 (1995).

Van Doren, M., Bailey, A.M., Esnayra, J., Ede, K. & Posakony, J.W. Negative regulation of proneural gene activity: hairy is a direct transcriptional repressor of achaete. *Genes & development* **8**, 2729-2742 (1994).

Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035-1041 (2010).

Vierbuchen, T. & Wernig, M. Direct lineage conversions: unnatural but useful? *Nature biotechnology* **29**, 892-907 (2011).

Vierbuchen, T. & Wernig, M. Molecular roadblocks for cellular reprogramming. *Molecular cell* **47**, 827-838 (2012).

Villares, R. & Cabrera, C.V. The achaete-scute gene complex of *D. melanogaster*: conserved domains in a subset of genes required for neurogenesis and their homology to myc. *Cell* **50**, 415-424 (1987).

Wapinski, O.L. *et al.* Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* **155**, 621-635 (2013).

Wernig, M. *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318-324 (2007).

Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448-452 (2003).

Wilson, M. & Koopman, P. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. *Current opinion in genetics & development* **12**, 441-446 (2002).

Yang, N., Ng, Y.H., Pang, Z.P., Sudhof, T.C. & Wernig, M. Induced neuronal cells: how to make and define a neuron. *Cell stem cell* **9**, 517-525 (2011).

Ying, F., Gong, Y., Sun, J.Y., Shen, J. & Han, X.D. [Primary culture, identification and functional study of rat Leydig cells]. *Zhonghua nan ke xue = National journal of andrology* **14**, 7-10 (2008).

Yoo, A.S. *et al.* MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* **476**, 228-231 (2011).

Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920 (2007).

Zaret, K.S. *et al.* Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. *Cold Spring Harbor symposia on quantitative biology* **73**, 119-126 (2008).

Zhao, Y. *et al.* Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell stem cell* **3**, 475-479 (2008).